MATHEMATISCHE ZEITSCHRIFT

63. BAND

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

K. KNOPP
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR A. E. INGHAM
H. KNESER W. MAGNUS O. PERRON
W. SÜSS H. WEYL†

63. BAND

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1955/56

Unveränderter Nachdruck 1971
Springer-Verlag Berlin Heidelberg New York

Inhalt des 63. Bandes

Seite

Issai Schur zum Gedächtnis 1

Behrend, F. A., A Contribution to the Theory of Magnitudes and the Foundations of Analysis 345

Bergman, S., Bounds for Analytic Functions in Domains with a Distinguished Boundary Surface 173

Bernays, P., Betrachtungen über das Vollständigkeitsaxiom und verwandte Axiome 219

Bochner, S., Green-Goursat Theorem 230

Brauer, A., On the Schnirelmann density of the sum of two sequences 529

Brauer, R., Zur Darstellungstheorie der Gruppen endlicher Ordnung 406

Dieudonné, J., Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$. III 53

Dinghas, A., Konvexitätseigenschaften von Mittelwerten harmonischer und verwandter Funktionen 109

Fekete, M., and G. Szegö, On algebraic equations with integral coefficients whose roots belong to a given point set 158

Fenchel, W., Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten 496

Freudenthal, H., Neuere Fassungen des Riemann-Helmholtz-Lieschen Raumproblems 374

Frucht, R., Zur Darstellung endlicher Abelscher Gruppen durch Kollineationen . . 145

Geiringer, H., Grenzlinien der Hodographentransformation 514

Grunsky, H., Eine funktionentheoretische Integralformel 320

Hahn, W., Über analytische Lösungen linearer Differential-Differenzengleichungen . 313

Hirsch, K. A., Über lokal-nilpotente Gruppen 290

Hopf, E., Die Harnacksche Ungleichung für positive harmonische Funktionen . . . 156

Jacobsthal, E., Über vertauschbare Polynome 243

Keller, O.-H., Eine Bemerkung zur Ausführung der körpertheoretischen Operationen in erträglich vielen Schritten 277

Knopp, K., Nörlund-Verfahren für Funktionen. (Zweite Mitteilung) 39

Kochendörffer, R., Über den Multiplikator einer Gruppe 507

Loewner, Ch., On totally positive matrices 338

Mordell, L. J., The Minimum of an Inhomogeneous Quadratic Polynomial in n Variables 525

Motzkin, Th. S., A proof of Hilbert's Nullstellensatz 341

Neumann, B. H., Groups with finite classes of conjugate subgroups 76

Ostrowski, A., Über Normen von Matrizen 2

Perron, O., Über die Abhängigkeit von Potenzsummen und einen Satz von Pólya . 19

Seite

Pólya, G., On the characteristic frequencies of a symmetric membrane 331

Rademacher, H., Zur Theorie der Dedekindschen Summen 445

Rado, R., Minimal Points of Convex Sets in Sequence Spaces 486

Rogosinski, W. W., On finite systems of linear equations with an infinity of unknowns 97

Rosenthal, A., On the continuity of functions of several variables 31

Rothe, E. H., Mapping degree in Banach spaces and spectral theory 195

Schulz, W., Über Reduzibilität bei gewissen Polynomen und das Tarry-Escottsche Problem 133

Siegel, C. L., Die Funktionalgleichungen einiger Dirichletscher Reihen 363

Skolem, Th., Einige Bemerkungen über die Auffindung der rationalen Punkte auf gewissen algebraischen Gebilden 295

Specht, W., Abschätzung der Wurzeln algebraischer Gleichungen. III 324

Stöhr, A., Zur Funktionentheorie im Raum der symmetrischen Matrizen 464

Taussky, O., Unimodular integral circulants 286

Wielandt, H., Primitive Permutationsgruppen vom Grad $2p$ 478

Issai Schur zum Gedächtnis.

Aus Anlaß der 80. Wiederkehr des Tages, an dem SCHUR in Mohilew am Dnjepr geboren wurde, vereinen sich Freunde und Schüler, um sein Andenken mit diesem Bande der Zeitschrift zu ehren, die er selbst begründet hat. Sie sind in alle Welt zerstreut durch die Katastrophe, in deren Verlauf SCHUR durch vorzeitige Emeritierung 1935 die Wirkungsstätte verlor, an der seine Vorlesungen drei Jahrzehnte lang Studenten für die Mathematik begeistert hatten. SCHUR konnte den furchtbarsten Folgen entkommen, indem er 1939 von Berlin nach Israel übersiedelte. Doch war es ihm nicht vergönnt, die wissenschaftliche Tätigkeit wieder aufzunehmen. Als er in Israel eintraf, war er bereits von dem Herzleiden gezeichnet, dem er am 10. Januar 1941, seinem 66. Geburtstag, in Tel Aviv erlag.

SCHUR hat, wie seine Antrittsrede in der Preußischen Akademie der Wissenschaften (1922) erkennen läßt, seine Aufgabe darin gesehen, die Algebra insbesondere auf dem Gebiet der linearen Transformationen zu fördern und in eine lebendige Wechselwirkung mit der Analysis zu bringen. Die Erfüllung dieser Aufgabe hat ihn als Mitbegründer der Darstellungstheorie in die Geschichte der Mathematik eingehen lassen. Doch enthalten auch seine zahlreichen auf andere Fragen gerichteten, stets zu vollendeter Form geschliffenen Arbeiten eine Fülle von Gedanken, die noch immer nicht ausgeschöpft ist. Möge dieser Band an SCHURs Gesamtwerk erinnern und von seiner Fruchtbarkeit zeugen.

Die Schriftleitung.

OSTROWSKI, A.
Math. Zeitschr. Bd. 63, S. 2—18 (1955)

Über Normen von Matrizen.

Dem Andenken an ISSAI SCHUR gewidmet.

Von

ALEXANDER OSTROWSKI.

Einleitung.

1. Im folgenden sollen die großen lateinischen Buchstaben $A, B, C, \ldots$ mit Ausnahme des § 3 quadratische Matrizen einer festen Ordnung n mit reellen oder komplexen Elementen $a_{\mu\nu}$ bzw. $b_{\mu\nu}, c_{\mu\nu}, \ldots$ bezeichnen. Werden andere Matrizen betrachtet, so wird dies jedesmal ausdrücklich angegeben werden. Die griechischen Buchstaben $\xi, \eta, \ldots$ bedeuten n-dimensionale (Kolonnen-) Vektoren mit den Komponenten $x_1, x_2, \ldots, x_n$ bzw. $y_1, y_2, \ldots, y_n, \ldots$.

2. Bei verschiedenartigsten Untersuchungen wurden bereits zur ungefähren Charakterisierung der „Größe" einer Matrix verschiedene Ausdrücke benutzt, die man als *Beträge, Schranken* oder auch *Normen* bezeichnet. In der Literatur sind vor allem sechs verschiedene solche Normen verwendet worden. Es sind dies in der weiter unten zu erläuternden Nomenklatur die drei *Hölderschen Normen*

$$(1) \qquad |A|_1 = \sum_{\mu,\nu} |a_{\mu\nu}|, \qquad |A|_2 = \left(\sum_{\mu,\nu} |a_{\mu\nu}|^2\right)^{\frac{1}{2}}, \qquad N^*(A) = |A|_\infty = \operatorname*{Max}_{\mu,\nu} |a_{\mu\nu}|,$$

sowie die zu den drei entsprechend definierten Vektornormen gehörenden symmetrischen *Schrankennormen*.

Es scheint nunmehr angebracht zu sein, eine systematische Untersuchung des Normbegriffes vorzunehmen unter Zugrundelegung geeigneter Postulate[1]). Sind auch, wie es sich herausstellt, für Limes- und Konvergenzbetrachtungen alle hier in Frage kommenden Normbegriffe im wesentlichen (für festes n) gleichberechtigt, so liefern sie andererseits in der Abschätzung der Eigenwerte oder bei der Diskussion des Fehlers in verschiedenen Iterationsverfahren sehr verschiedene Ergebnisse, so daß es wertvoll erscheint, über eine genügende Anzahl Normen zu verfügen. Dazu kommt noch, daß nicht alle Normen für Grenzübergänge mit $n \to \infty$ brauchbar sind. Es scheint, daß hierzu nur die symmetrischen Schrankennormen verwendet werden können.

[1]) Diese Untersuchung wurde für die sog. Schrankennormen in einer Note des Verfassers (OSTROWSKI [2]) begonnen und in allgemeineren Fällen von Herrn WE. GAUTSCHI in seiner Basler Dissertation ([2], [3]) fortgesetzt.

3. Wir definieren im folgenden eine *allgemeine Norm* $N(A)$ durch die drei Postulate:

I_N $N(A) > 0$ $(A \neq 0)$;

II_N $N(cA) = |c| N(A)$ für jede skalare Größe c;

III_N $N(A+B) \leq N(A) + N(B)$ [2]).

Aus II_N folgt offenbar für $c = 0$, daß $N(0) = 0$ ist.

4. Neben diesen Normen betrachten wir noch speziellere, sog. *multiplikative Normen* $M(A)$, die den folgenden vier Postulaten genügen (vgl. z.B. FADDEJEWA [1], wo aber I_N statt I_M benutzt wird, p. 60; NBSR p. 83):

I_M $M(A) \geq 0$ und für wenigstens ein A_0: $M(A_0) > 0$;

II_M $M(cA) = |c| M(A)$ für jede skalare Größe c;

III_M $M(A+B) \leq M(A) + M(B)$;

IV_M $M(AB) \leq M(A) M(B)$.

5. Wie man sieht, kommt hier das Postulat IV_M dazu; andererseits wird das Postulat I_M abgeschwächt. Wir beweisen nun im § 1, der der Diskussion der multiplikativen Normen gewidmet ist, den

Satz I. *Jede multiplikative Norm ist auch eine allgemeine Norm.*

Andererseits ergibt sich, daß $M(E)$, unter E die Einheitsmatrix verstanden, stets ≥ 1 ist. Offenbar ist keine ähnliche Relation bei allgemeinen Normen generell möglich, da ja mit $N(A)$ auch $cN(A)$ für jede positive Konstante c eine allgemeine Norm ist.

6. Wir bezeichnen in dieser Arbeit mit λ_A den maximalen absoluten Betrag der Fundamentalwurzeln der Matrix A. Eine besonders wichtige Eigenschaft multiplikativer Normen kommt zum Ausdruck im

Satz II. *Es gilt stets* $M(A) \geq \lambda_A$ [3]).

7. Für einen allgemeinen Vektor ξ wird eine allgemeine Norm $h(\xi)$ durch die Postulate definiert (vgl. z.B. FADDEJEWA [1], p. 59; NBSR p. 81):

I_V $h(\xi) > 0$ $(\xi \neq 0)$;

II_V $h(c\xi) = |c| h(\xi)$ für jede skalare Größe c;

III_V $h(\xi + \eta) \leq h(\xi) + h(\eta)$.

8. Sind nun $h(\xi)$ und $h_1(\xi)$ zwei derartige Vektornormen, so wird durch

$$S_{h,h_1}(A) = \operatorname{Sup}_{\xi \neq 0} \frac{h(A\xi)}{h_1(\xi)} \tag{2}$$

eine *Schrankennorm* von A definiert, die zu den Vektornormen $h(\xi)$ und $h_1(\xi)$ gehört. Ist insbesondere $h(\xi) = h_1(\xi)$, so haben wir es mit einer

[2]) Diese Postulate sind von GAUTSCHI [2], p. 375 als Postulate VI, VII und VIII benutzt worden.

[3]) Für zahlreiche Spezialfälle und auch für einige nicht multiplikative Normen hat GAUTSCHI ([1], [2], [3]) die Ungleichung des Satzes II bewiesen. Im allgemeinsten Falle einer multiplikativen Norm ist der Satz II auf einem anderen Wege als bei uns bewiesen bei FADDEJEWA [1], p. 65; NBSR p. 90.

symmetrischen Schrankennorm zu tun (vgl. z.B. FADDEJEWA [*1*], pp. 60, 61; NBSR pp. 84, 85):

$$S_h(A) = \operatorname{Sup}_{\xi \neq 0} \frac{h(A\xi)}{h(\xi)}. \tag{3}$$

Offenbar gilt für symmetrische Schrankennormen stets

$$S_h(E) = 1. \tag{4}$$

Übrigens folgt aus den Postulaten I_V bis III_V, daß eine Vektornorm stets eine *stetige Funktion der Vektorkomponenten* ist (vgl. den Stetigkeitsbeweis für allgemeine Normen in Nr. 20). Daher kann man in den Definitionen (2) und (3) das *Supremum* durch das *Maximum* ersetzen.

9. Aus der Definition (2) und den Postulaten I_V bis III_V folgt sofort, daß $S_{h,h_1}(A)$ den Postulaten I_N bis III_N genügt, also eine allgemeine Norm ist. Für die symmetrische Schrankennorm $S_h(A)$ folgt aber leicht, daß auch das Postulat IV_M erfüllt ist, so daß $S_h(A)$ stets eine *multiplikative Norm* ist. In der Tat gilt ja

$$\frac{h(AB\xi)}{h(\xi)} = \left(\frac{h(AB\xi)}{h(B\xi)}\right)\left(\frac{h(B\xi)}{h(\xi)}\right),$$

woraus IV_M sofort folgt. Im allgemeinsten Falle der nicht notwendig symmetrischen Schrankennormen wird in Nr. 19 als Formel (1.10) eine anders beschaffene „Multiplikationseigenschaft" aufgestellt.

10. Ist $N(A)$ eine allgemeine Matrizennorm, so läßt sich aus ihr eine Vektornorm $h(\xi)$ herleiten, indem man dem Vektor $\xi = (x_1, \ldots, x_n)$ eine $(n \times n)$-Matrix A_ξ zuordnet, die in der ersten Kolonne $x_1, x_2, \ldots, x_n$ hat und in den übrigen Kolonnen lauter Nullen. Dann kann man definieren

$$h_N(\xi) = N(A_\xi)^4). \tag{5}$$

Am Schlusse des § 1 beweisen wir den

Satz III. *Wird aus einer multiplikativen Norm $M(A)$ die zugehörige Vektornorm $h_M(\xi)$ hergeleitet und sodann die zu dieser Vektornorm gehörende symmetrische Schrankennorm $S_{h_M}(A)$ gebildet, so gilt stets*

$$M(A) \geq S_{h_M}(A). \tag{6}$$

11. Im § 2, der allgemeinen Normen $N(A)$ gewidmet ist, beweisen wir zuerst den

Satz IV. *Für zwei beliebige Normen $N(A)$ und $N_1(A)$ liegt der Quotient $\frac{N(A)}{N_1(A)}$ zwischen zwei positiven von A unabhängigen Konstanten.*

Es genügt hierzu zu beweisen, daß

$$p_1 < \frac{N(A)}{N^*(A)} < p_2 \tag{7}$$

---

4) Natürlich könnte man hier auch $N(A_\xi)$ so definieren, daß die Komponenten von A_ξ anstatt in der *ersten* Kolonne in der *k-ten* Kolonne von A_ξ stehen, und man erhält auf diese Weise mehrere aus $N(A)$ herleitbare Vektornormen, auf die unsere Betrachtungen genau so anwendbar bleiben.

gilt, wo $N^*(A)$ die in (1) definierte Norm ist und p_1 und p_2 von A unabhängige positive Zahlen sind.

Sodann beweisen wir eine Abschätzung von $N(A^t)$ für natürliche t, die für $t \to \infty$ die „richtige" Größenordnung von $N(A^t)$ liefert. Wir leiten nämlich her den

Satz V. *Es seien die Elementarteiler der Matrix A durch*

$$(\lambda - \lambda_1)^{m_1},\ (\lambda - \lambda_2)^{m_2},\ \ldots,\ (\lambda - \lambda_h)^{m_h} \tag{8}$$

gegeben, wobei die Anordnung so ist, daß

$$\lambda_A = |\lambda_1| = \cdots = |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_h| \tag{9}$$

und ferner

$$M \equiv m_1 = \cdots = m_s > m_{s+1} \geq \cdots \geq m_r \tag{10}$$

gilt. Dann besteht für jede allgemeine Norm $N(A)$ die Relation

$$p_3 < \frac{N(A^t)}{t^{M-1}\lambda_A^t} < p_4, \tag{11}$$

wo p_3 und p_4 zwei positive von t unabhängige und nur von $N(A)$ abhängende Zahlen sind. Gilt aber

$$\lambda_1 = \lambda_2 = \cdots = \lambda_s, \tag{12}$$

so konvergiert der Quotient in (11) *sogar gegen einen positiven endlichen Grenzwert*[5]).

Aus den Sätzen IV und V folgt dann sofort der

Satz VI. *Für jede allgemeine Norm gilt*

$$\sqrt[t]{N(A^t)} \to \lambda_A \qquad (t \to \infty)^6). \tag{13}$$

12. Im § 3 diskutieren wir die *Hölderschen Matrizennormen*

$$|A|_p = \Big(\sum_{\mu,\nu} |a_{\mu\nu}|^p\Big)^{1/p} \qquad (p \geq 1). \tag{14}$$

---

[5]) Die aus (11) folgende Abschätzung der allgemeinen Norm $N(A)$ *nach oben* hat in etwas schwächerer Form Herr GAUTSCHI ([1], [2]) bewiesen [dort ist M das Maximum sämtlicher Exponenten m_k in (8)]. Die zweite, unter der Annahme (12) geltende Behauptung des Satzes V hatte ich ursprünglich nur für die spezielle Norm $N^*(A)$ bewiesen. Erst durch Herrn GAUTSCHI bin ich in freundlicher Weise darauf aufmerksam gemacht worden, daß eine geringfügige Abänderung meines Beweises sie für jede allgemeine Norm liefert.

[6]) Die Relationen (7) und (13) sind unabhängig von mir auch von Herrn L. CARLESON bewiesen worden im Laufe einer für mich sehr anregenden Diskussion, für die ich ihm auch hier meinen herzlichsten Dank aussprechen möchte.

Es sei noch bemerkt, daß die Relation (13) für den allgemeinsten Fall von (symmetrischen und unsymmetrischen) Schrankennormen sich bei GAUTSCHI [2], p. 377, Formel (34), findet. Im Falle der symmetrischen Schrankennorm stellt sie einen Spezialfall eines allgemeinen Satzes der Theorie der linearen Transformationen im Hilbertschen Raum dar (vgl. RIESZ-NAGY [1], p. 421).

Bei dieser Diskussion werden wir die Annahme, daß die Matrizen $A, B, C, \ldots$ quadratische Matrizen sind, fallen lassen und sie für diesen ganzen Paragraphen nur als rechteckige Matrizen voraussetzen, wobei natürlich, damit das Produkt AB „existiert", vorausgesetzt werden muß, daß die Kolonnenanzahl von A gleich der Zeilenzahl von B ist. Wir beweisen dann den Satz VII, aus dem insbesondere folgen wird, daß *im Falle quadratischer Matrizen die Norm* $|A|_p$ *für* $1 \leq p \leq 2$ *eine multiplikative Norm ist*:

Satz VII. *Existiert das Matrizenprodukt* AB *und ist* $1 \leq p \leq 2$, *so gilt*[7])

$$|AB|_p \leq |A|_p |B|_p \qquad (1 \leq p \leq 2), \tag{15}$$

während (15) *für jedes* $p > 2$ *dann und nur dann allgemein*[8]) *gilt, wenn* A *aus einer einzigen Zeile und* B *aus einer einzigen Kolonne besteht.*

Für $p \geq 2$ gelten aber zwei der Hölderschen Ungleichung analoge Relationen.

Satz VIII. *Ist* $p \geq 2$, $\frac{1}{p} + \frac{1}{q} = 1$ *und existiert das Produkt* $C = AB$, *so gilt*

$$|AB|_p \leq |A|_p |B|_q, \tag{16}$$

$$|AB|_p \leq |A|_q |B|_p. \tag{17}$$

Dagegen gilt für jedes p *mit* $1 \leq p < 2$ *die Relation* (16) *dann und nur dann allgemein, wenn die Matrix* B *aus einer einzigen Kolonne besteht. Analog gilt* (17) *für jedes* p *mit* $1 \leq p < 2$ *dann und nur dann allgemein, wenn die Matrix* A *aus einer einzigen Zeile besteht.*

13. Im Falle einer *ungeraden* Anzahl von Faktoren gilt aber eine ähnliche Ungleichung für *alle* $p \geq 1$:

Satz IX. *Ist* $u > 1$ *ungerade, und existiert das Produkt* $A_1 A_2 \ldots A_u$, *so gilt für* $\frac{1}{p} + \frac{1}{q} = 1$

$$|A_1 A_2 \ldots A_u|_p \leq |A_1|_p |A_2|_q |A_3|_p \ldots |A_u|_p \qquad (p \geq 1). \tag{18}$$

Ist aber $g > 0$ *gerade und sind die Matrizen* $A_1, \ldots, A_g$ *die allgemeinsten Matrizen mit den vorgegebenen Zeilen- und Kolonnenanzahlen, für die das Produkt* $A_1 \ldots A_g$ *existiert, so gilt allgemein die Relation*

$$|A_1 A_2 \ldots A_g|_p \leq |A_1|_p |A_2|_q \ldots |A_{g-1}|_p |A_g|_q \tag{19}$$

dann und nur dann, wenn entweder $p \geq 2$ *oder die Kolonnenanzahl in* A_g *gleich* 1 *ist.*

Über die zu (17) analoge Ungleichung gilt der

Satz X. *Sind* $A_1, \ldots, A_l$ l *allgemeinste Matrizen mit vorgegebenen Zeilen- und Kolonnenanzahlen, deren Produkt existiert, so gilt im Falle eines ungeraden* l

$$|A_1 A_2 \ldots A_l|_p \leq |A_1|_q |A_2|_p \ldots |A_l|_q \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right) \tag{20}$$

---

[7]) Für $p = 2$ wurden diese Ungleichungen schon seit längerer Zeit verwendet (vgl. OSTROWSKI [1]).

[8]) Wir sagen hier (und später), daß eine solche Relation „allgemein" gilt, wenn sie für *alle* Matrizen mit festen Zeilen- und Kolonnenanzahlen gilt.

dann und nur dann allgemein, wenn entweder $p \geq 2$ oder die Zeilenanzahl in A_1 und die Kolonnenanzahl in A_l beide gleich 1 sind. Für ein gerades g gilt aber

$$(20^c) \qquad |A_1 A_2 \dots A_g|_p \leq |A_1|_q |A_2|_p \dots |A_g|_p \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right)$$

allgemein dann und nur dann, wenn entweder $p \geq 2$ oder die Zeilenanzahl in A_1 gleich 1 ist.

— Übrigens ergeben sich im Text die Beweise der Unmöglichkeitsbehauptungen der Sätze VII bis X, indem man sämtliche Elemente der Matrizen gleich 1 setzt.

Unsere Beweise der obigen Sätze machen fortgesetzt Gebrauch von der folgenden speziellen Norm der Matrix A:

$$(21) \qquad |A|_{q,p} = \left[\sum_{\nu} \left(\sum_{\mu} |a_{\mu\nu}|^q\right)^{p/q}\right]^{1/p} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right),$$

über die im § 3 verschiedene Hilfsformeln (3.9), (3.10), (3.11) und (3.12) hergeleitet werden. Diese spezielle Norm kann aber auch bei gewissen Abschätzungen von Matrizenpotenzreihen verwendet werden. Wir beweisen daher noch den

Satz XI. *Ist A eine quadratische Matrix, so gilt für jedes natürliche l*

$$(22) \qquad |A^l|_p \leq |A|_p |A|_{q,p}^{l-1} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

Sind allgemeiner $A_1, \dots, A_l$ rechteckige Matrizen, deren Produkt $A_1 \dots A_l$ existiert, so gilt

$$(23) \qquad |A_1 \dots A_l|_p \leq |A_1'|_{q,p} \dots |A_{k-1}'|_{q,p} |A_k|_p |A_{k+1}|_{q,p} \dots |A_l|_{q,p},$$

wo k auch gleich 1 oder l sein kann.

Endlich geben wir zum Schluß des § 3 verschiedene vereinzelte Formeln für die Schrankennormen, die zu den Hölderschen Vektornormen gehören.

§ 1. Multiplikative Normen.

14. Aus I_M und IV_M folgt für $A = A_0$ und $B = E$, wo E die Einheitsmatrix bedeutet, $M(A_0) \leq M(A_0) M(E)$,

$$(1.1) \qquad M(E) \geq 1.$$

15. Wir verstehen im folgenden unter $E_{\mu\nu}$ die Matrix, bei der das in der μ-ten Zeile und der ν-ten Kolonne stehende Element gleich 1 ist, während alle übrigen Elemente verschwinden. Dann gelten für jedes $A = (a_{\mu\nu})$ die allgemeinen Formeln

$$(1.2) \qquad A = \sum_{\sigma,\varkappa} a_{\sigma\varkappa} E_{\sigma\varkappa},$$

$$(1.3) \qquad E_{\mu\nu} A = \sum_{\varkappa} a_{\nu\varkappa} E_{\mu\varkappa}.$$

Um (1.3) aus (1.2) herzuleiten, genügt es zu beachten, daß allgemein, unter $\delta_{\sigma\nu}$ das Kroneckersche Symbol verstanden, gilt

$$(1.4) \qquad E_{\mu\nu} E_{\sigma\varkappa} = \delta_{\nu\sigma} E_{\mu\varkappa}.$$

Multiplizieren wir (1.3) rechtsseitig mit $E_{\sigma\mu}$ und beachten (1.4) für $\varkappa=\mu$, so folgt, wegen $E_{\mu\varkappa}E_{\sigma\mu}=\delta_{\varkappa\sigma}E_{\mu\mu}$,

$$E_{\mu\nu}AE_{\sigma\mu}=a_{\nu\sigma}E_{\mu\mu}. \tag{1.5}$$

Ersetzen wir in (1.5) A durch $E_{\nu\nu}$ und setzen $\sigma=\nu$, so ergibt sich

$$E_{\mu\nu}E_{\nu\nu}E_{\nu\mu}=E_{\mu\mu}. \tag{1.6}$$

16. Wegen $\sum_{\nu=1}^{n}E_{\nu\nu}=E$, wo E die Einheitsmatrix ist, folgt nach III_M und (1.1)

$$\sum_{\nu=1}^{n}M(E_{\nu\nu})>0.$$

Daher ist $M(E_{\mu\mu})$ für wenigstens ein μ positiv. Aus (1.6) folgt aber wegen IV_M

$$M(E_{\mu\mu})\leq M(E_{\mu\nu})\,M(E_{\nu\nu})\,M(E_{\nu\mu}), \tag{1.7}$$

so daß *alle* $M(E_{\mu\mu})$ positiv sind. Sodann folgt aber weiter aus (1.7), *daß* $M(E_{\mu\nu})$ *für alle* μ *und* ν *positiv ist.*

17. Wenden wir nun IV_M auf (1.5) an, so ergibt sich

$$|a_{\nu\sigma}|\,M(E_{\mu\mu})\leq M(E_{\mu\nu})\,M(E_{\sigma\mu})\,M(A). \tag{1.8}$$

Man setze nun

$$p=\operatorname*{Min}_{\mu,\nu}M(E_{\mu\nu}),\qquad P=\operatorname*{Max}_{\mu,\nu}M(E_{\mu\nu}).$$

Wählen wir dann ν und σ in (1.8) so, daß $|a_{\nu\sigma}|=\operatorname*{Max}_{\mu,\nu}|a_{\mu\nu}|=N^*(A)$ ist, so folgt schließlich für $A\neq 0$

$$M(A)\geq\frac{p}{P^2}N^*(A)>0,$$

so daß für $M(A)$ auch das Postulat I_N erfüllt ist. Damit ist der Satz I bewiesen.

18. Es sei nun ξ ein Eigenvektor von A, zu einer Fundamentalwurzel λ von A mit $|\lambda|=\lambda_A$ (vgl. Nr. 6) gehörend; dann gilt die Relation $A\xi=\lambda\xi$, die sich in den Bezeichnungen der Nr. 10 auch in der Form $AA_\xi=\lambda A_\xi$ schreiben läßt. Daraus folgt aber wegen II_M und IV_M

$$|\lambda|\,M(A_\xi)\leq M(A)\,M(A_\xi),$$

und da $M(A_\xi)>0$ ist, folgt daraus

$$M(A)\geq\lambda_A, \tag{1.9}$$

d.h. die Behauptung von Satz II.

19. Endlich ergibt sich der Satz III unmittelbar. In der Tat gilt nach (5)

$$\frac{h_M(A\xi)}{h_M(\xi)}=\frac{M(AA_\xi)}{M(A_\xi)}\leq\frac{M(A)\,M(A_\xi)}{M(A_\xi)}=M(A),$$

und (6) ergibt sich dann ohne weiteres aus der Definition (3).

In der Nr. 9 ist für allgemeine, durch die Formel (2) definierte Schrankennormen die Gültigkeit des Multiplikationspostulates IV_M nur im Falle der symmetrischen Schrankennormen bewiesen worden. Im Falle der allgemeinen

zu zwei Vektornormen h und h_1 gehörenden Schrankennormen gilt indessen eine etwas anders beschaffene „Multiplikationsrelation". Existiert nämlich das Produkt AB und sind h_1, h_2 und h_3 drei Vektornormen, so gilt

$$S_{h_1,h_3}(AB) \leq S_{h_1,h_2}(A)\, S_{h_2,h_3}(B), \tag{1.10}$$

sofern die Dimensionen von h_1, h_2, h_3 bzw. der Zeilenanzahl von A, der Zeilenanzahl von B und der Kolonnenanzahl von B gleich sind. In der Tat ist ja

$$\frac{h_1(AB\xi)}{h_3(\xi)} = \frac{h_1(AB\xi)}{h_2(B\xi)} \cdot \frac{h_2(B\xi)}{h_3(\xi)},$$

woraus (1.10) sofort folgt.

— Im Falle, daß die Normen h_1, h_2, h_3 als Quadratwurzeln aus positiven hermiteschen Formen in den Vektorkomponenten definiert werden, ist die Relation (1.10) von GAUTSCHI ([3], p. 4, Formel (15)) gegeben worden.

§ 2. Allgemeine Normen.

20. Wir beweisen nunmehr den Satz IV. Aus (1.2), II_N und III_N folgt

$$N(A) \leq \sum_{\mu,\nu} |a_{\mu\nu}| N(E_{\mu\nu}) \tag{2.1}$$

und daher

$$|N(A) - N(B)| \leq \sum_{\mu,\nu} |a_{\mu\nu} - b_{\mu\nu}|\, N(E_{\mu\nu}). \tag{2.2}$$

Daher ist $N(A)$ *stetig in den* $a_{\mu\nu}$, und zugleich gilt

$$N(A) \leq N^*(A) \sum_{\mu,\nu} N(E_{\mu\nu}) = p_2 N^*(A), \quad p_2 = \sum_{\mu,\nu} N(E_{\mu\nu}).$$

Andererseits existiert wegen der Stetigkeit von $N(A)$ auf der abgeschlossenen Menge $\operatorname{Max}_{\mu,\nu} |a_{\mu\nu}| = 1$ das

$$\operatorname{Min} N(A) = p_1$$

und ist positiv; daher gilt dann allgemein

$$N(A) \geq p_1 N^*(A),$$

und (7) ist bewiesen.

21. Wir kommen nunmehr zum Beweis des Satzes V. Es möge A durch die Matrix S auf die Jordansche Normalform[9]) entsprechend den Elementarteilern (8) transformiert werden:

$$T = SAS^{-1} = T_1 \dotplus T_2 \dotplus \ldots \dotplus T_k. \tag{2.3}$$

Hierbei *zerfällt* also T *vollständig* in die k quadratischen Matrizen $T_\varkappa$, wo $T_\varkappa$ eine Matrix der Ordnung $m_\varkappa$ ist, die zum Elementarteiler $(\lambda - \lambda_\varkappa)^{m_\varkappa}$ gehört. Jede dieser einzelnen Matrizen $T_\varkappa$ ist gleich

$$\lambda_\varkappa E_{m_\varkappa} + U_{m_\varkappa},$$

[9]) Vgl. z. B. SCHREIER und SPERNER [1], pp. 126—129, sowie namentlich TURNBULL and AITKEN [1], pp. 58—63.

wo $E_{m_\varkappa}$ die Einheitsmatrix der Ordnung $m_\varkappa$ ist, während $U_{m_\varkappa}$ eine quadratische Matrix der Ordnung $m_\varkappa$ ist, in der längs der Hauptdiagonalen und darunter lauter Nullen stehen, in der nächsten Superdiagonalen lauter Einsen und darüber wiederum lauter Nullen. Das Element $u_{\mu\nu}$ von $U_{m_\varkappa}$ ist also gegeben durch das Kroneckersche Symbol $\delta_{\nu,\mu+1}$.

Man beachte, daß wenn eine solche Matrix $U_{m_\varkappa}$ in die Potenz σ ($\sigma = 2, 3, \ldots$) erhoben wird, dann eine Matrix entsteht, die in der σ-ten Superdiagonalen lauter Einsen hat und sonst durchweg Nullen. Die Elemente $u_{\mu\nu}^{(\sigma)}$ von $U_{m_\varkappa}^\sigma$ sind daher gegeben durch $\delta_{\nu,\mu+\sigma}$. Insbesondere enthält $U_{m_\varkappa}^{m_\varkappa-1}$ nur eine Einheit in der rechten oberen Ecke und sonst durchweg Nullen, während $U_{m_\varkappa}^\sigma$ für $\sigma \geq m_\varkappa$ verschwindet.

22. Daher gilt

$$(\lambda_\varkappa E_{m_\varkappa} + U_{m_\varkappa})^t = \lambda_\varkappa^t \left[E_{m_\varkappa} + \binom{t}{1}\lambda_\varkappa^{-1} U_{m_\varkappa} + \binom{t}{2}\lambda_\varkappa^{-2} U_{m_\varkappa}^2 + \cdots + \binom{t}{m_\varkappa - 1}\lambda_\varkappa^{-m_\varkappa+1} U_{m_\varkappa}^{m_\varkappa-1}\right].$$

Daraus folgt für den maximalen absoluten Betrag eines Elementes der Matrix links die für hinreichend große t geltende Relation

$$N^*(T_\varkappa^t) = \binom{t}{m_\varkappa - 1}|\lambda_\varkappa|^{t-m_\varkappa+1}.$$

Darüber hinaus gilt für $t \to \infty$ offenbar

$$T_\varkappa^t = \binom{t}{m_\varkappa - 1}\lambda_\varkappa^{t-m_\varkappa+1} E_{1m_\varkappa}^{(\varkappa)} + O(t^{m_\varkappa-2}|\lambda_\varkappa|^t),$$

wo $E_{1m_\varkappa}^{(\varkappa)}$ eine quadratische Matrix der Ordnung $m_\varkappa$ ist, die in der rechten oberen Ecke eine Eins enthält und sonst lauter Nullen.

23. Nunmehr folgt aus (2.3) für hinreichend große t

$$N^*(T^t) = \operatorname*{Max}_\varkappa \binom{t}{m_\varkappa - 1}|\lambda_\varkappa|^{t-m_\varkappa+1} = \lambda_A^t \operatorname*{Max}_{\varkappa \leq s}\binom{t}{m_\varkappa - 1}\lambda_A^{-m_\varkappa+1} = \binom{t}{M-1}\lambda_A^{t-M+1},$$

$$N^*(T^t) = \binom{t}{M-1}\lambda_A^{t-M+1} \tag{2.4}$$

und, $\lambda_\nu/\lambda_A = \varepsilon_\nu$, ($\nu = 1, 2, \ldots, s$) gesetzt,

$$T^t = \binom{t}{M-1}\lambda_A^{t-M+1}[E_{1M}\varepsilon_1^{t-M+1} + E_{M+1,2M}\varepsilon_2^{t-M+1} + \cdots + E_{(s-1)M+1,sM}\varepsilon_s^{t-M+1}] + O(t^{M-2}\lambda_A^t),$$

wo nunmehr die Symbole $E_{\mu\nu}$ die in der Nr. 15 festgelegte Bedeutung haben. Daraus folgt im Falle, daß die Relation (12) gilt, da ja dann alle ε_ν einander gleich sind,

$$T^t = \binom{t}{M-1}\lambda_1^{t-M+1}\sum_{\sigma=1}^{s} E_{(\sigma-1)M+1,\sigma M} + O(t^{M-2}\lambda_A^t). \tag{2.5}$$

24. Wegen (2.3) gilt $T^t = SA^tS^{-1}$, $A^t = S^{-1}T^tS$ für jedes natürliche t und daraus folgt für

$$(2.6)\qquad \begin{cases} S=(s_{\mu\nu}), \quad S^{-1}=(s'_{\mu\nu}), \quad \sigma=\sum\limits_{\mu,\nu}|s_{\mu\nu}|\sum\limits_{\mu,\nu}|s'_{\mu\nu}|: \\ N^*(T^t)\leq\sigma N^*(A^t), \quad N^*(A^t)\leq\sigma N^*(T^t) \end{cases}$$

und daher wegen (2.4)

$$(2.7)\qquad \frac{1}{\sigma}\leq\frac{N^*(A^t)}{\binom{t}{M-1}\lambda_A^{t-M+1}}\leq\sigma.$$

womit wegen (7) auch (11) bewiesen ist.

25. Andererseits folgt aus (2.5) für $t\to\infty$

$$\frac{A}{\binom{t}{M-1}\lambda_1^{t-M+1}}=C+O\left(\frac{1}{t}\right), \qquad C=\sum_{\sigma=1}^{s}S^{-1}E_{(\sigma-1)M+1,\sigma M}S.$$

Daher gilt wegen II_N in Nr. 3 für $t\to\infty$

$$\frac{N(A^t)}{\binom{t}{M-1}\lambda_A^{t-M+1}}\to N(C).$$

Wäre nun $N(C)=0$, so wäre $N(A^t)=o(t^{M-1}\lambda_A^t)$, entgegen (2.7). Daher folgt die zweite Behauptung des Satzes V unmittelbar[10]).

§ 3. Höldersche Normen.

26. Von nun an sollen die Matrizen $A, B, \ldots$ nur als rechteckige Matrizen vorausgesetzt werden. Auch für solche Matrizen, z.B. Matrizen vom $(m\times n)$-Typus, d.h. mit m Zeilen und n Kolonnen, soll der Normbegriff durch die Postulate I_N bis III_N der Nr. 3 definiert werden. Eine solche Norm kann z.B. in sehr allgemeiner Weise aus gegebenen Vektornormen für n- und m-dimensionale Vektoren gebildet werden. Es seien $h_1(\xi), \ldots, h_n(\xi)$ n gleiche oder verschiedene Vektornormen für m-dimensionale Vektoren ξ und es sei $h(\eta)$ eine Vektornorm für n-dimensionale Vektoren η. $h(\eta)$ soll ferner die Eigenschaft haben, daß sie für nicht negative Komponenten $y_1, \ldots, y_n$ von η eine nicht abnehmende Funktion dieser Komponenten ist. Bezeichnen wir dann mit ξ_ν den ν-ten Kolonnenvektor der Matrix A, so daß also

$$\xi_\nu=(a_{1\nu}, a_{2\nu}, \ldots, a_{m\nu})$$

ist, so setze man

$$(3.1)\qquad H(A)=h[h_1(\xi_1), h_2(\xi_2), \ldots, h_n(\xi_n)].$$

Damit ist eine Norm von A definiert.

---

[10]) Wir hätten anstatt der Jordanschen Normalform auch z.B. die explizite Darstellung von A^t, wie sie sich bei H. Schmidt ([1], p. 549, Formel (9)) findet, benutzen können, doch wäre hierzu erst eine weitere Diskussion der in diese Formel eingehenden Vektoren notwendig.

27. In der Tat, daß die Postulate I_N und II_N der Nr. 3 erfüllt sind, ist unmittelbar klar. Bei III_N ergibt sich dies aber wie folgt: Es sei A^0 eine weitere $(m \times n)$-Matrix und es seien ihre Kolonnenvektoren $\xi_1^0, \ldots, \xi_n^0$. Dann ergibt sich wegen (3.1) unter Benutzung der Postulate der Nr. 7:

$$H(A+A^0) = h[h_1(\xi_1+\xi_1^0), \ldots, h_n(\xi_n+\xi_n^0)] \leq$$
$$\leq h[h_1(\xi_1)+h_1(\xi_1^0), \ldots, h_n(\xi_n)+h_n(\xi_n^0)] \leq$$
$$\leq h[h_1(\xi_1), \ldots, h_n(\xi_n)] + h[h_1(\xi_1^0), \ldots, h_n(\xi_n^0)] = H(A)+H(A^0).$$

28. Definieren wir z.B. allgemein für einen k-dimensionalen Vektor ξ und für ein p, $1 \leq p \leq \infty$,

$$|\xi|_p = \Big(\sum_\varkappa |x_\varkappa|^p\Big)^{1/p} \qquad (1 \leq p \leq \infty), \tag{3.2}$$

wobei natürlich für $p=\infty$, $|\xi|_\infty = \operatorname{Max}_\varkappa |x_\varkappa|$ ist, so ist $|\xi|_p$ eine Norm im Sinne der Postulate der Nr. 7. Dabei folgt insbesondere das Postulat III_V aus der Minkowskischen Ungleichung.

Setzen wir nun insbesondere in (3.1) $h_\nu(\xi) = |\xi|_{p_1}$ und $h(\eta) = |\eta|_{p_2}$, so erhalten wir die Norm

$$|A|_{p_1,p_2} = \Big[\sum_\nu \Big(\sum_\mu |a_{\mu\nu}|^{p_1}\Big)^{p_2/p_1}\Big]^{1/p_2}, \tag{3.3}$$

wo p_1 und p_2 zwischen 1 und ∞ liegen. Andererseits ergibt sich, wenn (3.2) auf den Fall $k = m\,n$ angewandt wird, daß der für eine Matrix A durch (14) definierte Ausdruck für $1 \leq p \leq \infty$ eine Norm ist.

Man beachte, daß während $|A|_p$ beim Übergang zur transponierten Matrix unverändert bleibt, $|A'|_p = |A|_p$, dies bei der Norm (3.3) im allgemeinen nicht der Fall ist. — Ist die Kolonnenanzahl in A gleich 1, so gilt, wie man sich sofort überzeugt,

$$|A|_{p_1,p_2} = |A|_{p_1}.$$

29. Mit Hilfe von (3.2) läßt sich die Höldersche Ungleichung so formulieren: Gilt

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad 1 \leq p \leq \infty, \tag{3.4}$$

so gilt für das innere Produkt

$$(\xi, \eta) = \sum_{\nu=1}^{n} x_\nu \bar{y}_\nu \tag{3.5}$$

der Vektoren ξ, η der Dimension n, die Relation

$$|(\xi, \eta)| \leq |\xi|_p |\eta|_q. \tag{3.6}$$

Andererseits gilt bekanntlich, daß der Ausdruck (3.2) *mit zunehmendem $p \geq 1$ monoton abnimmt.* (Diese Tatsache ist zuerst von A. PRINGSHEIM angegeben worden (vgl. HARDY-LITTLEWOOD-PÓLYA [1], p. 28).)

Ist nun $p \leq 2$, so ist wegen (3.4) $q \geq 2$, $p \leq q$, so daß

$$|\xi|_q \leq |\xi|_p \qquad (p \leq 2 \leq q) \tag{3.7}$$

gilt. Daher folgt nunmehr aus (3.6), wenn (3.7) auf η angewendet wird,

$$|(\xi,\eta)| \leq |\xi|_p |\eta|_p \qquad (1 \leq p \leq 2). \tag{3.8}$$

30. Wir beweisen nun einen Hilfssatz über den Ausdruck (3.3) unter der Annahme, daß $p_1 = q$ und $p_2 = p$ durch die Relation (3.4) verknüpft sind.

Hilfssatz. *Gilt* (3.4), *so gilt für zwei Matrizen* A, B, *für die das Produkt* AB *existiert*:

$$|AB|_{q,p} \leq |A|_q |B|_p. \tag{3.9}$$

Ist aber darüber hinaus $p \geq 2$, *so gilt*

$$|A|_{q,p} \leq |A|_q \qquad (p \geq 2). \tag{3.10}$$

31. *Beweis.* Es sei A eine $(m \times k)$- und B eine $(k \times n)$-Matrix. Dann ist $AB = C = (c_{\mu\nu})$ eine $(m \times n)$-Matrix und es gilt wegen $c_{\mu\nu} = \sum_\varkappa a_{\mu\varkappa} b_{\varkappa\nu}$ nach (3.6), wenn dort p mit q vertauscht wird und beide Seiten in die q-te Potenz erhoben werden,

$$|c_{\mu\nu}|^q \leq \Big(\sum_\varkappa |a_{\mu\varkappa}|^q\Big)\Big(\sum_\varkappa |b_{\varkappa\nu}|^p\Big)^{q/p}$$

und daher, wenn nach μ summiert wird, wegen (14),

$$\sum_\mu |c_{\mu\nu}|^q \leq |A|_q^q \Big(\sum_\varkappa |b_{\varkappa\nu}|^p\Big)^{q/p},$$

$$\Big(\sum_\mu |c_{\mu\nu}|^q\Big)^{p/q} \leq |A|_q^p \sum_\varkappa |b_{\varkappa\nu}|^p,$$

und wenn hier über ν summiert wird, wegen (3.3) und (14),

$$|AB|_{q,p}^p \leq |A|_q^p |B|_p^p,$$

womit (3.9) bewiesen ist.

32. Andererseits läßt sich, wenn

$$y_\nu = \Big(\sum_\mu |a_{\mu\nu}|^q\Big)^{1/q},$$

gesetzt wird, $|A|_{q,p}$ schreiben als

$$|A|_{q,p} = \Big(\sum_\nu y_\nu^p\Big)^{1/p}$$

und dies ist, wenn $p \geq 2 \geq q$ ist, nach (3.7), wenn dort p mit q vertauscht wird, höchstens gleich

$$\Big(\sum_\nu y_\nu^q\Big)^{1/q} = \Big(\sum_{\nu,\mu} |a_{\mu\nu}|^q\Big)^{1/q} = A_q,$$

womit (3.10) bewiesen ist.

33. Um nun den Satz VII zu beweisen, nehmen wir an, daß A eine $(m \times k)$- und B eine $(k \times n)$-Matrix ist. Ist dann $c_{\mu\nu}$ das allgemeine Element von $C = AB$, so gilt $c_{\mu\nu} = \sum_\varkappa a_{\mu\varkappa} b_{\varkappa\nu}$ und daher wegen (3.8) für $p \leq 2$

$$|c_{\mu\nu}|^p \leq \sum_\varkappa |a_{\mu\varkappa}|^p \sum_\varkappa |b_{\varkappa\nu}|^p.$$

Summieren wir hier über μ und ν, so folgt

$$|C|_p^p \leq |A|_p^p |B|_p^p,$$

woraus (15) unmittelbar folgt.

34. Ist aber $p>2$, so wählen wir als Matrizen A und B die Matrizen, deren sämtliche Elemente gleich 1 sind. Dann ist $c_{\mu\nu}=k$ und

$$|A|_p = (m\,k)^{1/p}, \quad |B|_p = (n\,k)^{1/p}, \quad C_p = k\,(m\,n)^{1/p}.$$

Hier liefe die Relation (15) hinaus auf

$$(m\,n)^{1/p}\,k \leq m^{1/p}\,k^{1/p}\,n^{1/p}\,k^{1/p},$$

was für $p>2$ unmöglich ist, solange $k>1$ ist. Für $k=1$ dagegen ist (15) offenbar für alle $p\geq 1$ richtig, sogar mit dem Gleichheitszeichen. Damit ist der Satz VII bewiesen.

35. Es sei wiederum A eine $(m\times k)$- und B eine $(k\times n)$-Matrix. Dann gilt nach (3.6) für $c_{\mu\nu}=\sum_{\varkappa} a_{\mu\varkappa} b_{\varkappa\nu}$:

$$|c_{\mu\nu}|^p \leq \sum_{\varkappa} |a_{\mu\varkappa}|^p \Big(\sum_{\varkappa} |b_{\varkappa\nu}|^q\Big)^{p/q}$$

und daher, wenn nach μ und ν summiert wird, wegen (3.3)

$$|AB|_p^p \leq |A|_p^p \sum_{\nu} \Big(\sum_{\varkappa} |b_{\varkappa\nu}|^q\Big)^{p/q} = |A|_p^p\,|B|_{q,p}^p,$$

$$|AB|_p \leq |A|_p\,|B|_{q,p} \qquad \Big(\frac{1}{p}+\frac{1}{q}=1\Big). \tag{3.11}$$

Andererseits folgt wiederum aus (3.6), wenn dort p mit q vertauscht wird:

$$|c_{\mu\nu}|^p \leq \sum_{\varkappa} |b_{\varkappa\nu}|^p \Big(\sum_{\varkappa} |a_{\mu\varkappa}|^q\Big)^{p/q},$$

und daraus durch Summation nach μ und ν

$$|AB|_p^p \leq |B|_p^p \sum_{\mu} \Big(\sum_{\varkappa} |a_{\mu\varkappa}|^q\Big)^{p/q}.$$

Der zweite Faktor rechts ist aber gleich $|A'|_{q,p}^p$, unter A' die zu A gehörende transponierte Matrix verstanden, so daß wir schließlich erhalten

$$|AB|_p \leq |B|_p\,|A'|_{q,p}. \tag{3.12}$$

36. Ist nun $p\geq 2$, so folgt aus (3.11) und (3.10), auf B angewandt, die Relation (16). Aus (3.12) und (3.10) folgt aber

$$|AB|_p \leq |B|_p\,|A'|_q = |B|_t\,|A|_q,$$

womit auch (17) bewiesen ist.

37. Ist jetzt $p<2$, so spezialisiere man A und B wiederum zu Matrizen, deren sämtliche Elemente gleich 1 sind. Dann gelten die Relationen

$$|A|_p = (m\,k)^{1/p}, \quad |B|_q = (n\,k)^{1/q}, \quad |C|_p = k\,(m\,n)^{1/p}.$$

Dann liefe die Ungleichung (16) hinaus auf

$$(m\,n)^{1/p} \leq m^{1/p}\,k^{1/p}\,n^{1/q}\,k^{1/q},$$

d.h. wegen (3.4) auf $n^{1/p} \leq n^{1/q}$, was für $n > 1$, $p < 2$ unmöglich ist. Ebenso ist (17) für $m > 1$, $p < 2$ unmöglich.

38. Ist aber $n = 1$, so folgt aus der Definition (3.3), daß $|B|_{q,p} = |B|_q$, so daß (16) für alle $p \geq 1$ aus (3.11) folgt. Ebenso reduziert sich für $m = 1$, $|A'|_{q,p}$ auf $|A'|_q = |A|_q$, so daß dann (17) aus (3.12) für alle $p \geq 1$ folgt. Damit ist der Satz VIII bewiesen.

39. Um die Relation (18) des Satzes IX zu beweisen, beachte man, daß aus (3.11) folgt

$$|A_1 A_2 \dots A_u|_p \leq |A_1|_p \, |A_2 \dots A_u|_{q,p} \tag{3.13}$$

und daher, wenn auf den zweiten Faktor rechts (3.9) angewendet wird,

$$|A_1 A_2 \dots A_u|_p \leq |A_1|_p \, |A_2|_q \, |A_3 \dots A_u|_p, \tag{3.14}$$

woraus (18) durch vollständige Induktion hervorgeht.

40. Soll nun die Relation (19) allgemein gelten und ist A_1 eine $(m \times k_1)$-, A_2 eine $(k_1 \times k_2)$-, ..., A_g eine $(k_{g-1} \times n)$-Matrix, so spezialisiere man alle diese Matrizen in der Weise, daß ihre sämtlichen Elemente gleich 1 gesetzt werden. Dann ist das Produkt eine $(m \times n)$-Matrix, deren sämtliche Elemente gleich $k_1 k_2 \dots k_{g-1}$ sind. Daher wird die linke Seite von (19) gleich

$$(m n)^{1/p} k_1 k_2 \dots k_{g-1},$$

während das Produkt rechts in (19) zu

$$(m k_1)^{1/p} (k_1 k_2)^{1/q} \dots (k_{g-2} k_{g-1})^{1/p} (k_{g-1} n)^{1/q} = m^{1/p} n^{1/q} k_1 \dots k_{g-1}$$

wird. Aus (19) folgt aber dann $n^{1/p} \leq n^{1/q}$, was nur für $n = 1$ oder $p \geq 2 \geq q$ gelten kann.

Um aber einzusehen, daß in den beiden letzten Fällen (19) gilt, beachte man, daß im Falle eines geraden g genau wie oben aus (3.13) und (3.14) durch vollständige Induktion die Formel

$$|A_1 A_2 \dots A_g|_p \leq |A_1|_p \, |A_2|_q \dots |A_{g-1}|_p \, |A_g|_{q,p} \tag{3.15}$$

folgt. Für $p \geq 2$ folgt aber dann (19) aus (3.10) und für $n = 1$ daraus, daß dann wegen (3.3) der letzte Faktor in (3.15) gleich $|A_g|_q$ wird. Damit ist Satz IX bewiesen.

41. Soll nun die Relation (20) des Satzes X allgemein gelten, so spezialisiere man die Matrizen $A_1, \dots, A_l$ wiederum derart, daß ihre sämtlichen Elemente gleich 1 werden, und nehme die zugehörigen Zeilen- und Kolonnenanzahlen, wie in der Nr. 40, bzw. gleich $m, k_1, k_2, \dots, k_{l-1}, n$. Dann ist $A_1 \dots A_l$ eine $(m \times n)$-Matrix, deren sämtliche Elemente gleich $k_1 \dots k_{l-1}$ sind, so daß der Ausdruck links in (20) zu $(m n)^{1/p} k_1 \dots k_{l-1}$ wird. Das Produkt rechts in (20) wird im Falle eines ungeraden l, wie in der Nr. 40, zu $(m k_1)^{1/q} (k_1 k_2)^{1/p} \dots (k_{l-1} n)^{1/q}$, so daß hier (20) auf

$$(m n)^{1/p} \leq (m n)^{1/q}$$

hinausläuft, und wir entweder $m=n=1$ oder $p\geq 2\geq q$ erhalten. Für ein gerades l dagegen wird das Produkt rechts in (20) zu $(m\,k_1)^{1/q}(k_1k_2)^{1/p}\dots(k_{l-1}n)^{1/p}$ und (20) läuft auf

$$(m\,n)^{1/p}=m^{1/q}n^{1/p}$$

hinaus, so daß entweder $m=1$ oder $p\geq 2\geq q$ ist.

42. Ist andererseits $p\geq 2$, so folgt aus (17)

$$|A_1\dots A_l|_p\leq|A_1|_q|A_2\dots A_l|_p$$

und (20) bzw. (20°) folgt hier aus (18) oder (19), da ja auch die Relation (19) für $p\geq 2$ gilt. Ist aber $m=1$, so folgt aus (3.12)

$$|A_1\dots A_l|_p\leq|A_2\dots A_l|_p|A_1'|_{q,p}.$$

Der letzte Faktor ist aber gleich $|A_1'|_q=|A_1|_q$, da die Kolonnenzahl m in A_1' den Wert 1 hat. Daher folgt

$$|A_1\dots A_l|_p\leq|A_1|_q|A_2\dots A_l|_p \qquad (m=1). \tag{3.16}$$

Ist nun g gerade, also die Faktorenzahl in $A_2\dots A_g$ ungerade, so folgt (20°) aus (3.16) wegen (18). Ist dagegen l ungerade, so haben wir auch zugleich $n=1$ vorausgesetzt. In diesem Falle ist aber auf den zweiten Faktor rechts in (3.16) die Ungleichung (19) anzuwenden, so daß (20) folgt. Damit ist der Satz X bewiesen.

43. Zum Beweise des Satzes XI genügt es offenbar (23) zu beweisen, da (22) daraus für $A_1=A_2=\dots=A_l$, $k=1$ ohne weiteres folgt. Nun folgt aber aus (3.12)

$$|A_1\dots A_l|_p\leq|A_1'|_{q,p}|A_2\dots A_l|_p\leq\dots\leq|A_1'|_{q,p}\dots|A_{k-1}'|_{q,p}|A_kA_{k+1}\dots A_l|_p.$$

Für den letzten Faktor rechts folgt durch wiederholte Anwendung von (3.11)

$$|A_k\dots A_l|_p\leq|A_k\dots A_{l-1}|_p|A_l|_{q,p}\leq\dots\leq|A_k\dots A_{l-2}|_p|A_{l-1}|_{q,p}|A_l|_{q,p}\leq\dots\leq$$
$$\leq|A_k|_p|A_{k+1}|_{q,p}\dots|A_l|_{q,p}.$$

Damit ist Satz XI bewiesen.

44. Wir wollen noch einige Angaben über diejenigen Schrankennormen [vgl. die Definition in (2)] hinzufügen, die zu den durch (3.2) definierten Hölderschen Vektornormen gehören. Wir setzen

$$S_{p_1,p_2}(A)=\operatorname{Sup}\frac{|A\xi|_{p_1}}{|\xi|_{p_2}} \qquad (1\leq p_1,p_2\leq\infty) \tag{3.17}$$

und ferner

$$S_p(A)=S_{p,p}(A) \qquad (1\leq p\leq\infty). \tag{3.18}$$

Aus dem in Nr. 29 erwähnten monotonen Abnehmen von (3.2) folgt, daß $S_{p_1,p_2}(A)$ eine *monoton abnehmende* Funktion von p_1 und eine *monoton zunehmende* Funktion von p_2 ist. Von den symmetrischen Schrankennormen

(3.18) lassen sich namentlich die $p=1$ und $p=\infty$ entsprechenden sehr einfach darstellen:

$$S_1(A) = \operatorname*{Max}_{\nu} \sum_{\mu} |a_{\mu\nu}|, \tag{3.19}$$

$$S_\infty(A) = \operatorname*{Max}_{\mu} \sum_{\nu} |a_{\mu\nu}|. \tag{3.20}$$

Diese beiden sehr leicht zu beweisenden Formeln werden häufig in der Iterationstheorie der Matrizen verwendet[11]).

45. Es ist von Interesse, den Wert der Norm (3.17) für $A=E$, wo E eine n-dimensionale Einheitsmatrix ist, zu bestimmen. Hierzu ist es offenbar notwendig das Extremalproblem zu lösen

$$\operatorname{Max}\left(\sum_{\nu=1}^{n} x_\nu^{p_1}\right)^{1/p_1} = ? \qquad \left(\sum_{\nu=1}^{n} x_\nu^{p_2} = 1\right). \tag{3.21}$$

Ist $p_1 \geq p_2$, so folgt aus der oben erwähnten Monotonieeigenschaft von $S_{p_1,p_2}(A)$, daß $S_{p_1,p_2}(E) \leq S_{p_2,p_2}(E) = 1$ ist, und die Schranke 1 wird erreicht für $x_1=1$, $x_2=x_3=\cdots=x_n=0$. Für $p_1<p_2$ liefert die Lagrangesche Multiplikatorenmethode sofort das Resultat, daß wenn keines der x_ν im Extremalpunkt verschwindet, $x_1=x_2=\cdots=x_n=n^{-1/p_2}$ und dafür $|\xi|_{p_1} = n^{1/p_1-1/p_2} > 1$ ist. Wären aber einige der $x_\nu=0$, so hätte man hier n durch eine kleinere Zahl zu ersetzen, so daß sich schließlich ergibt:

$$S_{p_1,p_2}(E) = \begin{cases} 1 & (p_1 \geq p_2) \\ n^{\frac{1}{p_1}-\frac{1}{p_2}} & (p_1 \leq p_2). \end{cases} \tag{3.22}$$

46. Existiert das Matrizenprodukt AB, so folgt, wenn (18) auf das Zählerprodukt angewandt wird:

$$S_{p_1,p_2}(AB) = \operatorname{Sup} \frac{|AB\xi|_{p_1}}{|\xi|_{p_2}} \leq |A|_{p_1} |B|_{q_1} S_{p_1,p_2}(E) \qquad \left(\frac{1}{p_1}+\frac{1}{q_1}=1\right)$$

und daher insbesondere wegen (3.22)

$$S_{p_1,p_2}(AB) \leq |A|_{p_1} |B|_{q_1} \qquad (p_1 \geq p_2). \tag{3.24}$$

Ferner folgt, wenn die Sätze VII und VIII auf $|A\xi|_{p_1}$ angewandt werden,

$$S_{p_1,p_2}(A) \leq |A|_{p_1} S_{p_1,p_2}(E) \qquad (p_1 \leq 2), \tag{3.25}$$

$$S_{p_1,p_2}(A) \leq |A|_{p_1} S_{q_1,p_2}(E) \qquad (1 \leq p_1 \leq \infty), \tag{3.26}$$

$$S_{p_1,p_2}(A) \leq |A|_{q_1} S_{p_1,p_2}(E) \qquad (p_1 \geq 2). \tag{3.27}$$

Es sei endlich noch vermerkt, daß

$$S_{\infty,1}(A) = N^*(A) = \operatorname*{Max}_{\mu,\nu} |a_{\mu\nu}| \tag{3.28}$$

[11]) Vgl. z. B. FADDEJEWA [1], pp. 62, 63; NBSR pp. 85—87.

ist. In der Tat gilt

$$\frac{\operatorname{Max}_{\mu} \left| \sum_{\nu} a_{\mu\nu} x_\nu \right|}{\sum_{\nu} |x_\nu|} \leq \operatorname{Max}_{\mu,\nu} |a_{\mu\nu}|,$$

und die Schranke rechts wird offenbar erreicht, wenn ein geeignetes der x_ν gleich 1 und alle übrigen gleich 0 gesetzt werden.

Literatur.

AITKEN, A. C., and H. W. TURNBULL: [*1*] An Introduction to the Theory of Canonical Matrices. London and Glasgow, Blackie and Son 1950. — FADDEJEWA, W. N.: [*1*] Numerische Methoden der linearen Algebra. Moskau 1950 (russisch). Eine Übersetzung des für uns vor allem in Frage kommenden 1. Kapitels des Buches ist veröffentlicht worden als National Bureau of Standards Report 1644, Computational methods of linear algebra, Ch. 1, Basic material from linear algebra, pp. V + 1—93 (Translator: C. D. Benster, Editor: G. E. Forsythe). Im Text werden wir die Zitate auf diese Übersetzung mit NBSR charakterisieren. — GAUTSCHI, WE.: [*1*] The asymptotic behaviour of powers of matrices. Duke Math. J. **20**, 127—140 (1953). — [*2*] The asymptotic behaviour of powers of matrices. II. Duke Math. J. **20**, 375—379 (1953). — [*3*] Bounds of matrices with regard to an Hermitian metric. Comp. Math. **12**, 1—16 (1954). — HARDY, LITTLEWOOD, PÓLYA: [*1*] Inequalities. Cambridge 1934. — OSTROWSKI, AS: [*1*] Sur la détermination des bornes inférieures pour une classe des déterminants. Bull. Sci. Math. (2) **61**, 1—14 (1937). — [*2*] Un nouveau théorème d'existence pour les systèmes d'équations. C. R. Acad. Sci. Paris **232**, 786—788 (1951). — RIESZ, F., et B. NAGY: [*1*] Lecons d'analyse fonctionelle, deuxième edition. Budapest 1953. — SCHMIDT, H.: [*1*] Über Wurzelapproximation nach Euler und Fixgebilde linearer Transformationen. Math. Z. **52**, 547—556 (1950). — SCHREIER, O., u. E. SPERNER: [*1*] Einführung in die analytische Geometrie und Algebra, Bd. II. Leipzig 1935.

Mathematische Anstalt der Universität Basel (Schweiz).

(Eingegangen am 24. Dezember 1954.)

PERRON, O.
Math. Zeitschr. Bd. 63, S. 19—30 (1955)

Über die Abhängigkeit von Potenzsummen und einen Satz von Pólya.

ISSAI SCHUR zum Gedächtnis.

Von

OSKAR PERRON.

Zu den Lieblingsgebieten von I. SCHUR haben die algebraischen Gleichungen und die Eigenschaften ihrer Wurzeln gehört. Noch in der letzten Arbeit, deren Erscheinen in Band 4 der Compositio Mathematica (1937) er erleben durfte, hat er sich mit den Potenzsummen befaßt. Wenn er auch zumeist wohl arithmetische Fragen im Auge hatte, so dürften doch die nachstehenden Betrachtungen, bei denen es sich um Potenzsummen von *n Variablen (Unbestimmten)* handelt, in einem Band, der seinem Andenken gewidmet ist, nicht fehl am Ort sein.

§ 1. Abhängigkeit von Potenzsummen.

Zwischen $n+1$ Formen (d.h. homogenen Polynomen) $f_0, f_1, \ldots, f_n$ von n Variablen $x_1, \ldots, x_n$ besteht stets eine Abhängigkeit

$$\sum A_{\lambda_0 \lambda_1 \ldots \lambda_n} f_0^{\lambda_0} f_1^{\lambda_1} \cdots f_n^{\lambda_n} = 0,$$

und zwar, wenn mit m_ν der Grad von f_ν und mit d der größte gemeinsame Teiler von $m_0, m_1, \ldots, m_n$ bezeichnet wird, eine Abhängigkeit von der Höhe (dem Gewicht)

$$H = \frac{m_0 m_1 \cdots m_n}{d},$$

d.h. jedes Glied der Summe ist in bezug auf die x_ν vom Grad

$$m_0 \lambda_0 + m_1 \lambda_1 + \cdots + m_n \lambda_n = H.$$

Wenn die f_ν *allgemeine* Formen sind, d.h., wenn die Koeffizienten Unbestimmte sind, dann besteht zudem keine Abhängigkeit von kleinerer Höhe[1]). Zwischen spezialisierten Formen kann aber sehr wohl eine Abhängigkeit kleinerer Höhe bestehen. Im folgenden sollen speziell gewisse symmetrische Formen betrachtet werden.

[1]) Dieser Doppelsatz ist bewiesen in meiner Arbeit: „Über die Abhängigkeit von Polynomen", Sitzungsberichte Bayer. Akad. Wiss., Math.-nat. Klasse **1950**, 117—130. Einen zweiten Beweis hat Herr HEINZ ORSINGER in seiner Dissertation gegeben: „Zur Konstruktion von Trägheitsformen als Koeffizienten algebraischer Gleichungen." Math. Nachr. **5**, 355—370 (1951).

Trivial ist zunächst folgendes: Sind

$$\begin{aligned} a_1 &= -(x_1 + x_2 + \cdots + x_n), \\ a_2 &= +(x_1 x_2 + x_1 x_3 + \cdots + x_{n-1} x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_n &= (-1)^n x_1 x_2 \cdots x_n \end{aligned}$$

die symmetrischen Grundfunktionen und ist f eine symmetrische Form vom Grad m, so ist f bekanntlich als Polynom von $a_1, a_2, \ldots, a_n$ darstellbar:

$$f = \sum B_{\lambda_1 \lambda_2 \ldots \lambda_n} a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n},$$

wobei jedes Glied der Summe vom Gewicht $\lambda_1 + 2\lambda_2 + \cdots + n\lambda_n = m$ ist. Diese Formel bedeutet bereits eine Abhängigkeit der Höhe m zwischen $a_1, \ldots, a_n$ und f, und eine Abhängigkeit von kleinerer Höhe besteht offenbar nicht. Insbesondere kann man die Potenzsummen

$$s_\nu = x_1^\nu + x_2^\nu + \cdots + x_n^\nu$$

so darstellen, und wenn etwa $\nu < n$ ist, s_ν schon als Polynom von $a_1, \ldots, a_\nu$. Die Darstellung ergibt sich leicht in Determinantenform aus den bekannten NEWTONschen Formeln, die ebenso auch umgekehrt a_ν als Polynom von $s_1, \ldots, s_\nu$ darzustellen gestatten. Infolgedessen ist jede symmetrische Form f vom Grad m auch in der Gestalt

$$f = \sum C_{\lambda_1 \lambda_2 \ldots \lambda_n} s_1^{\lambda_1} s_2^{\lambda_2} \ldots s_n^{\lambda_n}$$

darstellbar, wobei offenbar wieder $\lambda_1 + 2\lambda_2 + \cdots + n\lambda_n = m$ ist. Diese Formel bedeutet eine Abhängigkeit der Höhe m zwischen $s_1, \ldots, s_n$ und f, und eine Abhängigkeit kleinerer Höhe besteht wieder nicht. Insbesondere die Abhängigkeit zwischen $s_1, \ldots, s_n, s_{n+1}$ ergibt sich, indem man $a_1, \ldots, a_n$ aus den ersten $n+1$ NEWTONschen Formeln eliminiert, in der Gestalt:

$$\begin{vmatrix} s_1 & 1 & & & \\ s_2 & s_1 & 2 & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_n & s_{n-1} & s_{n-2} & \cdots s_1 & n \\ s_{n+1} & s_n & s_{n-1} & \cdots s_2 & s_1 \end{vmatrix} = 0,$$

wo die leeren Stellen durch Nullen zu besetzen sind.

Ganz anders und durchaus nicht mehr trivial ist die Sache, wenn nur Potenzsummen mit ungeraden Exponenten betrachtet werden. Die ersten n dieser Potenzsummen $s_1, s_3, s_5, \ldots, s_{2n-1}$ sind augenscheinlich voneinander unabhängig, denn ihre Funktionaldeterminante ist eine nicht verschwindende VANDERMONDEsche Determinante. Dagegen besteht zwischen $s_1, s_3, s_5, \ldots, s_{2n+1}$ nach obigem Satz eine Abhängigkeit der Höhe $1 \cdot 3 \cdot 5 \cdots (2n+1)$. Aber das ist nicht die kleinstmögliche Höhe, vielmehr werden wir folgendes beweisen:

Satz 1. *Zwischen den* $n+1$ *Potenzsummen*

$$s_\nu = x_1^\nu + x_2^\nu + \cdots + x_n^\nu \qquad (\nu = 1, 3, 5, \ldots, 2n+1)$$

besteht genau eine Abhängigkeit der Höhe $H = (n+1)(n+2)/2$, *die zudem in bezug auf* s_{2n+1} *linear ist, so daß* s_{2n+1} *sich rational durch* $s_1, s_3, \ldots, s_{2n-1}$ *ausdrücken läßt. Es besteht aber keine Abhängigkeit von kleinerer Höhe.*

Außerdem werden wir die Abhängigkeit der Höhe H auch formelmäßig angeben können.

Zunächst ist klar, daß es *höchstens eine* Abhängigkeit der Höhe H geben kann. Denn gäbe es zwei verschiedene, so könnte man daraus s_{2n+1} eliminieren und hätte eine Abhängigkeit zwischen $s_1, s_3, \ldots, s_{2n-1}$, während diese n Formen doch voneinander unabhängig sind. Nunmehr zeigen wir, daß eine Abhängigkeit von kleinerer Höhe als H

$$\text{(I)} \qquad \begin{cases} \sum D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}} s_1^{\lambda_1} s_3^{\lambda_3} \ldots s_{2n+1}^{\lambda_{2n+1}} = 0, \\ \lambda_1 + 3\lambda_3 + 5\lambda_5 + \cdots + (2n+1)\lambda_{2n+1} = H' < H \end{cases}$$

jedenfalls nicht besteht. Nehmen wir nämlich an, sie bestünde doch, so bilden wir nach Einführung von zwei weiteren Variablen x_{n+1}, x_{n+2} die Potenzsummen

$$S_\nu = x_1^\nu + x_2^\nu + \cdots + x_{n+1}^\nu + x_{n+2}^\nu = s_\nu + x_{n+1}^\nu + x_{n+2}^\nu.$$

Bei der Spezialisierung $x_{n+2} = -x_{n+1}$ wird $S_\nu = s_\nu$ (für ungerade ν). Bilden wir daher die Form

$$\varphi(x_1, \ldots, x_n, x_{n+1}, x_{n+2}) = \sum D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}} S_1^{\lambda_1} S_3^{\lambda_3} \ldots S_{2n+1}^{\lambda_{2n+1}}$$

mit den obigen Koeffizienten $D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}}$, so geht diese bei der Spezialisierung $x_{n+1} + x_{n+2} = 0$ über in das Polynom (I), verschwindet also. Infolgedessen ist sie durch $x_{n+1} + x_{n+2}$ teilbar. Ebenso ist sie aber als symmetrische Form durch jede Summe $x_\nu + x_\mu$, also auch durch das Produkt

$$\prod_{1 \le \nu < \mu \le n+2} (x_\nu + x_\mu)$$

teilbar. Dieses Produkt ist aber vom Grad $(n+1)(n+2)/2 = H$, während die Form φ nur vom kleineren Grad H' ist, also nicht teilbar sein kann, ohne identisch zu verschwinden. Daher ist

$$\sum D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}} S_1^{\lambda_1} S_3^{\lambda_3} \ldots S_{2n+1}^{\lambda_{2n+1}} = 0.$$

Das besagt nun, daß zwischen den $n+1$ Potenzsummen $S_1, S_3, \ldots, S_{2n+1}$ von $n+2$ Variablen $x_1, \ldots, x_{n+1}\, x_{n+2}$ eine Abhängigkeit besteht, während sie doch voneinander unabhängig sind (wie oben mit $n-2$ an Stelle von n festgestellt wurde). Die Annahme der Abhängigkeit (I) führt also zu einem Widerspruch.

Um nun die Abhängigkeit der Höhe H zu finden, setzen wir

$$\text{(1)} \qquad F(t) = (1 - x_1 t) \ldots (1 - x_n t) = 1 + a_1 t + a_2 t^2 + \cdots + a_n t^n.$$

Dann ist[2])

$$\frac{F'(t)}{F(t)} = \sum_{\nu=1}^{n} \frac{-x_\nu}{1-x_\nu t} = -\sum_{\nu=1}^{n} (x_\nu + x_\nu^2 t + x_\nu^3 t^2 + \cdots)$$
$$= -(s_1 + s_2 t + s_3 t^2 + \cdots).$$

Wenn man hier t durch $-t$ ersetzt, ergibt sich durch Addition

$$\frac{F'(t)}{F(t)} + \frac{F'(-t)}{F(-t)} = -2(s_1 + s_3 t^2 + s_5 t^4 + \cdots). \tag{2}$$

Multipliziert man mit den Nennern herauf und setzt dann die beiderseits entstehenden Koeffizienten von $t^0, t^2, t^4, \ldots, t^{2n}$ einander gleich, so erhält man $n+1$ in $s_1, s_3, \ldots, s_{2n+1}$ lineare Gleichungen, die aber in bezug auf die a_ν nicht linear sind, so daß sich die a_ν nicht ohne weiteres eliminieren lassen. Zu linearen Gleichungen für die a_ν gelangt man nach einer Idee von PÓLYA auf folgende Weise[3]):

Setzt man

$$\frac{F(-t)}{F(t)} = 1 + w_1 t + w_2 t^2 + w_3 t^3 + \cdots, \tag{3}$$

so ergibt sich durch logarithmische Ableitung mit Rücksicht auf (2):

$$\left\{ \begin{aligned} \frac{w_1 + 2w_2 t + 3w_3 t^2 + \cdots}{1 + w_1 t + w_2 t^2 + \cdots} &= -\left(\frac{F'(t)}{F(t)'} + \frac{F'(-t)}{F(-t)}\right) \\ &= 2(s_1 + s_3 t^2 + s_5 t^4 + \cdots). \end{aligned} \right. \tag{4}$$

Wenn man hier mit dem linken Nenner heraufmultipliziert und dann gleiche Potenzen von t links und rechts vergleicht, erhält man zur Berechnung der w_ν die Rekursionsformeln

$$\left\{ \begin{aligned} &w_1 = 2s_1, \qquad 2w_2 = 2s_1 w_1, \\ &(2\nu-1)\, w_{2\nu-1} = 2s_1 w_{2\nu-2} + 2s_3 w_{2\nu-4} + \cdots + 2s_{2\nu-3} w_2 + 2s_{2\nu-1}, \\ &2\nu\, w_{2\nu} = 2s_1 w_{2\nu-1} + 2s_3 w_{2\nu-3} + \cdots + 2s_{2\nu-3} w_3 + 2s_{2\nu-1} w_1, \end{aligned} \right. \tag{5}$$

wonach klar ist, daß die w_ν Polynome von $s_1, s_3, s_5, \ldots$ sind, und zwar w_ν vom Gewicht ν. Für $\nu \leq 7$ ist z. B.

$$w_1 = 2s_1, \quad w_2 = 2s_1^2, \quad w_3 = \tfrac{1}{3}(4s_1^3 + 2s_3), \quad w_4 = \tfrac{1}{3}(2s_1^4 + 4s_1 s_3),$$
$$w_5 = \tfrac{1}{15}(4s_1^5 + 20s_1^2 s_3 + 6s_5), \quad w_6 = \tfrac{1}{45}(4s_1^6 + 40s_1^3 s_3 + 10s_3^2 + 36s_1 s_5),$$
$$w_7 = 3\tfrac{1}{15}(8s_1^7 + 140s_1^4 s_3 + 140s_1 s_3^2 + 252s_1^2 s_5 + 90s_7).$$

Aus (5) ergibt sich weiter durch vollständige Induktion, daß $w_{2\nu-1}$ das Glied $\frac{2s_{2\nu-1}}{2\nu-1}$ enthält und $w_{2\nu}$ für $\nu > 1$ das Glied $\frac{4s_1 s_{2\nu-1}}{2\nu-1}$; ferner enthält w_ν sowohl für gerades als ungerades ν das Glied $\frac{2^\nu s_1^\nu}{\nu!}$, welches lediglich für $\nu = 1$ mit dem zuvor angegebenen zusammenfällt.

[2]) Mit den hier auftretenden Potenzreihen, deren Koeffizienten keine Zahlen sind, so daß Konvergenz nicht in Frage kommt, läßt sich ganz formal rechnen. Man könnte aber (unbequemer) auch mit endlichen Summen und Restglied rechnen, analog wie ich es z. B. in Band 1 meiner „Algebra", S. 160, 2. und 3. Aufl. S. 151 f. aus didaktischen Gründen getan habe.

[3]) PÓLYA, G.: Remarques sur un problème d'algèbre étudié par Laguerre. J. de math. pures et appl. **31**, 37—47 (1952).

Nun folgt aus (3) durch Multiplikation mit $F(t)$:

$$(1 + w_1 t + w_2 t^2 + \cdots)(1 + a_1 t + a_2 t^2 + \cdots a_n t^n) = 1 - a_1 t + a_2 t^2 - \cdots + (-1)^n a_n t^n.$$

Vergleicht man hier beiderseits die Koeffizienten von $t^1, t^3, t^5, \ldots, t^{2n+1}$, so kommt

$$(6)\quad \begin{cases} w_1 + 2a_1 = 0, \\ w_3 + w_2 a_1 + w_1 a_2 + 2a_3 = 0, \\ w_5 + w_4 a_1 + w_3 a_2 + w_2 a_3 + w_1 a_4 + 2a_5 = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_{2n+1} + w_{2n} a_1 + w_{2n-1} a_2 + \cdots + w_1 a_{2n} + 2a_{2n+1} = 0, \end{cases}$$

wobei $a_\nu = 0$ für $\nu > n$ zu setzen ist. Aus diesen $n+1$ Gleichungen kann man $a_1, a_2, \ldots, a_n$ eliminieren und erhält

$$(7)\quad \begin{vmatrix} w_1 & 2 & & & & \\ w_3 & w_2 & w_1 & 2 & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-1} & w_{2n-2} & w_{2n-3} & \cdots & \cdots & w_{n-1} \\ w_{2n+1} & w_{2n} & w_{2n-1} & \cdots & \cdots & w_{n+1} \end{vmatrix} = 0,$$

wo die leeren Stellen durch Nullen zu besetzen sind. Diese Formel stellt nun, da die w_ν Polynome von $s_1, s_3, s_5, \ldots$ vom Gewicht ν sind, eine Abhängigkeit zwischen den $n+1$ Potenzsummen $s_1, s_3, \ldots s_{2n+1}$ dar, und zwar, wie man am einfachsten aus dem Diagonalglied $w_1 w_2 w_3 \ldots w_{n+1}$ der Determinante (7) feststellt, von der Höhe

$$1 + 2 + 3 + \cdots + (n+1) = \tfrac{1}{2}(n+1)(n+2) = H.$$

Sie ist also die gesuchte Abhängigkeit der Höhe H, vorausgesetzt allerdings, daß die Determinante (7) als Polynom der s_ν nicht etwa identisch verschwindet. Das tut sie aber nicht. Denn sie enthält, da in w_{2n+1} das Glied $\frac{2s_{2n+1}}{2n+1}$ steckt und die übrigen in der Determinante (7) vorkommenden w_ν von s_{2n+1} unabhängig sind (ihr Index, also ihr Gewicht, ist ja kleiner als $2n+1$), jedenfalls den Bestandteil

$$(8)\quad \begin{cases} (-1)^n \frac{2s_{2n+1}}{2n+1} \begin{vmatrix} 2 & & & & \\ w_2 & w_1 & 2 & & \\ w_4 & w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-2} & w_{2n-3} & \cdots & \cdots & w_{n-1} \end{vmatrix} \\ \quad = (-1)^n \frac{4s_{2n+1}}{2n+1} \begin{vmatrix} w_1 & 2 & & & \\ w_3 & w_2 & w_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & \cdots & \cdots & w_{n-1} \end{vmatrix}. \end{cases}$$

Hier enthält nun die rechts stehende Determinante entsprechend den Bestandteil

$$(-1)^{n-2}\frac{4s_{2n-3}}{2n-3}\begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots \\ w_{2n-7} & w_{2n-8} & \cdots & w_{n-3} \end{vmatrix}.$$

Durch Fortsetzung dieses Verfahrens erkennt man, daß die Determinante (7) für gerades $n=2r$ das Glied

$$\frac{4s_{4r+1}}{4r+1}\cdot\frac{4s_{4r-3}}{4r-3}\cdots\frac{4s_5}{5}\cdot\frac{2s_1}{1}$$

enthält[4]) und für ungerades $n=2r-1$ das Glied

$$(-1)^r\frac{4s_{4r-1}}{4r-1}\cdot\frac{4s_{4r-5}}{4r-5}\cdots\frac{4s_7}{7}\cdot\frac{4s_3}{3}.$$

Sie verschwindet also als Polynom der s_ν nicht identisch. Unsere Überlegung zeigt zugleich, daß die Determinante (7) den Bestandteil (8) enthält und daß im übrigen s_{2n+1} nicht mehr darin vorkommt. Die Abhängigkeit (7) ist also, wie in Satz 1 behauptet, in bezug auf s_{2n+1} linear.

Es ergibt sich beispielsweise

für $n=1$: $$\begin{vmatrix} w_1 & 2 \\ w_3 & w_2 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $-3/4$: $s_1^3 - s_3 = 0$,

für $n=2$: $$\begin{vmatrix} w_1 & 2 & \\ w_3 & w_2 & w_1 \\ w_5 & w_4 & w_3 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $3\cdot 3\cdot 5/8$:

$$s_1^6 - 5s_1^3 s_3 - 5s_3^2 + 9s_1 s_5 = 0,$$

für $n=3$: $$\begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ w_5 & w_4 & w_3 & w_2 \\ w_7 & w_6 & w_5 & w_4 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $3\cdot 3\cdot 3\cdot 5\cdot 5\cdot 7/16$:

$$s_1^{10} - 15s_1^7 s_3 + 63s_1^5 s_5 - 225s_1^3 s_7 + 315s_1^2 s_3 s_5 - 175s_1 s_3^3 + 225s_3 s_7 - 189s_5^2 = 0.$$

[4]) Kein Druckfehler! Im letzten Faktor heißt es $2s_1$, nicht $4s_1$. Denn bei Fortsetzung des Verfahrens stößt man zuletzt auf die einreihige Determinante $|w_1|$, zu der kein Faktor 2 mehr hinzutritt, und es ist $w_1 = 2s_1$.

§ 2. Ein Satz von Pólya.

Läßt man von den $n+1$ Gleichungen (6) die letzte weg, so kann man die ersten n nach $a_1, a_2, \ldots, a_n$ auflösen. Die dabei auftretende Nennerdeterminante ist, nach den Elementen der ersten Zeile entwickelt, gleich $2W_n$, wobei

$$(9) \qquad W_n = \begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & \cdots & w_{n-1} \end{vmatrix} \qquad (W_1 = 1,\ W_2 = w_1)$$

ist. Daß W_n als Polynom der s_ν nicht identisch verschwindet, kann zunächst genau so gezeigt werden wie oben bei der Determinante (7)[5]). Wir werden es aber gleich noch auf eine zweite Art beweisen, indem wir für W_n als Polynom der x_ν einen anderen Ausdruck herleiten. Wird die Variable x_n zu $-x_{n-1}$ spezialisiert[6]), dann geht s_ν für ungerades ν über in $x_1^\nu + x_2^\nu + \cdots + x_{n-2}^\nu$. Somit geht W_n genau in die Determinante (7) mit $n-2$ an Stelle von n über [man beachte dabei die Fußnote [5])]. Da aber (7) als Polynom der x_ν verschwindet, so besagt das, wenn man die für n gewonnene Erkenntnis auf $n-2$ anwendet, daß W_n bei der Spezialisierung $x_{n-1} + x_n = 0$ verschwindet. Infolgedessen ist W_n durch $x_{n-1} + x_n$ teilbar. Ebenso ist W_n als symmetrische Form auch durch jeden Faktor $x_\nu + x_\mu$ und folglich durch das Produkt

$$(10) \qquad P_n = \prod_{1 \le \nu < \mu \le n} (x_\nu + x_\mu)$$

teilbar. Nun ist P_n vom Grad $n(n-1)/2$ und die Determinante W_n ist homogen und, wie man sofort aus dem Diagonalglied $w_1 w_2 \ldots w_{n-1}$ ersieht, ebenfalls vom Grad $n(n-1)/2$. Somit ist

$$(11) \qquad W_n = K_n P_n,$$

wo K_n eine Konstante ist, die wir jetzt durch geeignete Spezialisierung der x_ν berechnen wollen.

Im Fall eines ungeraden $n = 2r-1$ spezialisieren wir die x_ν zu den $(2r-1)$-ten Einheitswurzeln, also

$$x_\nu = e^{2\nu\varphi i} \quad (\nu = 1, 2, \ldots, 2r-1), \qquad \text{wobei} \qquad \varphi = \frac{\pi}{2r-1}.$$

Dann ist

$$F(t) = \prod_{\nu=1}^{2r-1} (1 - e^{2\nu\varphi i} t) = 1 - t^{2r-1},$$

$$\frac{F(-t)}{F(t)} = \frac{1 + t^{2r-1}}{1 - t^{2r-1}} = \frac{1 + t^n}{1 - t^n} = 1 + 2t^n + 2t^{2n} + \cdots.$$

---

[5]) Man beachte übrigens, daß n auch die Anzahl der Variablen x_ν bezeichnet, so daß die Determinante (7), die aus der Determinante (9) formal dadurch hervorgeht, daß n durch $n+2$ ersetzt wird, keineswegs mit W_{n+2} zu bezeichnen ist. Denn in (7) und (9) sind die w_ν Formen von n Variablen, während sie bei W_{n+2} Formen von $n+2$ Variablen wären.

[6]) Das geht nur für $n > 2$. Aber die Fälle $n = 1$ und $n = 2$ erledigen sich von selbst. Es ist $W_1 = 1$ und $W_2 = w_1 = 2s_1 = 2(x_1 + x_2)$.

Also ist $w_n = 2$, während alle anderen in W_n auftretenden w_ν verschwinden. Die Determinante W_n hat $n-1 = 2r-2$ Reihen und sieht jetzt so aus:

$$\left.\begin{vmatrix} 0 & 2 & & & & & \\ 0 & 0 & 0 & 2 & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 0 & 2 \end{vmatrix}\right\} r-1$$

$$\left.\begin{vmatrix} 2 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 2 & 0 \end{vmatrix}\right\} r-1$$

In jeder Zeile und in jeder Spalte steht ein Zweier, und zwar haben die Zweier mit den Zeilennummern $1, 2, \dots, 2r-2$ der Reihe nach die Spaltennummern

$$2, 4, 6, \dots, 2r-2, 1, 3, 5, \dots, 2r-3.$$

Die Anzahl der Inversionen ist

$$1 + 2 + 3 + \cdots + (r-1) = r(r-1)/2,$$

und folglich ist

$$W_{2r-1} = (-1)^{r(r-1)/2} \cdot 2^{2r-2}.$$

Zur Berechnung von P_{2r-1} notieren wir zunächst die aus der Definition von $F(t)$ für $t = -e^{-2\mu\varphi i}$ hervorgehende Formel

$$\prod_{\nu=1}^{2r-1} (1 + e^{2(\nu-\mu)\varphi i}) = 1 + 1 = 2,$$

aus der einerseits nach Multiplikation mit $(e^{2\mu\varphi i})^{2r-1} = 1$, anderseits für $\mu = 0$ und nach Division durch $1 + e^{2(2r-1)\varphi i} = 2$ folgt:

$$(12) \qquad \prod_{\nu=1}^{2r-1} (e^{2\mu\varphi i} + e^{2\nu\varphi i}) = 2, \quad \prod_{\nu=1}^{2r-2} (1 + e^{2\nu\varphi i}) = 1.$$

Nunmehr ist

$$P_{2r-1}^2 \cdot \prod_{\nu=1}^{2r-1} (x_\nu + x_\nu) = \prod_{\mu=1}^{2r-1} \prod_{\nu=1}^{2r-1} (x_\mu + x_\nu)$$

$$= \prod_{\mu=1}^{2r-1} \prod_{\nu=1}^{2r-1} (e^{2\mu\varphi i} + e^{2\nu\varphi i}) = \prod_{\mu=1}^{2r-1} 2 = 2^{2r-1},$$

und da das links stehende einfache Produkt gleich

$$\prod_{\nu=1}^{2r-1} (2e^{2\nu\varphi i}) = 2^{2r-1} \cdot e^{2r(2r-1)\varphi i} = 2^{2r-1}$$

ist, ergibt sich zunächst $P_{2r-1}^2 = 1$. Die Berechnung von P_{2r-1} selbst gestaltet sich nun so: Es ist

$$P_{2r-1} = \prod_{\nu<\mu} (e^{2\nu\varphi i} + e^{2\mu\varphi i}) = \prod_{\nu<\mu} e^{2\nu\varphi i} \cdot \prod_{\nu<\mu} (1 + e^{2(\mu-\nu)\varphi i})$$
$$= e^{2a\varphi i} \cdot \prod_{\mu=2}^{2r-1} \prod_{\nu=1}^{\mu-1} (1 + e^{2(\mu-\nu)\varphi i}),$$

wobei es auf die ganze Zahl a nicht ankommen wird, ebenso auch später bei den ganzen Zahlen b, c. Setzt man

$$\prod_{\nu=1}^{\mu-1} (1 + e^{2(\mu-\nu)\varphi i}) = \eta_\mu,$$

so ist also

$$P_{2r-1} = e^{2a\varphi i} \prod_{\mu=2}^{2r-1} \eta_\mu = e^{2a\varphi i} \cdot \eta_{2r-1} (\eta_{2r-2}\eta_2)(\eta_{2r-3}\eta_3) \cdots (\eta_{r+1}\eta_{r-1})\eta_r.$$

Nun ist nach der zweiten Formel (12): $\eta_{2r-1} = 1$; ferner

$$\eta_{r+k}\eta_{r-k} = \prod_{\nu=1}^{r+k-1} (1 + e^{2(r+k-\nu)\varphi i}) \cdot \prod_{\nu=1}^{r-k-1} (1 + e^{2(r-k-\nu)\varphi i})$$
$$= \prod_{\nu=1}^{r+k-1} e^{2(r+k-\nu)\varphi i} \cdot \prod_{\nu=1}^{r+k-1} (e^{2(r-1-k+\nu)\varphi i} + 1) \cdot \prod_{\nu=1}^{r-k-1} (1 + e^{2(r-k-\nu)\varphi i})$$
$$= e^{2b\varphi i} \cdot \prod_{\nu=1}^{2r-2} (1 + e^{2\nu\varphi i}) = e^{2b\varphi i} \quad \text{(nach (12))}.$$

Außerdem ist

$$\eta_r = \prod_{\nu=1}^{r-1} (1 + e^{2(r-\nu)\varphi i}) = \prod_{\nu=1}^{r-1} e^{(r-\nu)\varphi i} \cdot \prod_{\nu=1}^{r-1} (e^{-(r-\nu)\varphi i} + e^{(r-\nu)\varphi i})$$
$$= e^{r(r-1)/2 \cdot \varphi i} \cdot \prod_{\nu=1}^{r-1} (2\cos(r-\nu)\varphi).$$

Setzt man das alles in den zuletzt gefundenen Ausdruck für P_{2r-1} ein, so ergibt sich

$$P_{2r-1} = e^{2c\varphi i} \cdot e^{r(r-1)/2 \cdot \varphi i} \cdot \prod_{\nu=1}^{r-1} (2\cos(r-\nu)\varphi).$$

Da die Cosinus offenbar positiv sind und da wir bereits wissen, daß $P_{2r-1}^2 = 1$ ist, kommt also schließlich

$$P_{2r-1} = e^{[2c + r(r-1)/2]\varphi i}.$$

Wegen $P_{2r-1}^2 = 1$ muß $2c + r(r-1)/2$ durch $2r-1$ teilbar sein. Setzt man demgemäß

$$2c + r(r-1)/2 = k(2r-1), \quad \text{so daß} \quad k \equiv r(r-1)/2 \pmod 2$$

ist, so ergibt sich schließlich

$$P_{2r-1} = e^{k(2r-1)\varphi i} = e^{k\pi i} = (-1)^k = (-1)^{r(r-1)/2}.$$

Für die Konstante K_n in (11) findet man also den Wert

$$K_n = K_{2r-1} = \frac{W_{2r-1}}{P_{2r-1}} = 2^{2r-2} = 2^{n-1}.$$

Im Fall eines geraden $n = 2r$ spezialisieren wir $x_{2r} = 0$ und

$$x_\nu = e^{2\nu\varphi i} \quad (\nu = 1, 2, \ldots, 2r-1), \quad \text{wobei} \quad \varphi = \frac{\pi}{2r-1}.$$

Dann ist wieder $F(t) = 1 - t^{2r-1}$ und

$$\frac{F(-t)}{F(t)} = \frac{1 + t^{2r-1}}{1 - t^{2r-1}} = \frac{1 + t^{n-1}}{1 - t^{n-1}} = 1 + 2t^{n-1} + 2t^{2n-2} + \cdots.$$

Daher ist diesmal $w_{n-1} = 2$ und alle anderen in W_n auftretenden w_ν verschwinden. Die Determinante W_n hat $n - 1 = 2r - 1$ Reihen und sieht so aus:

$$\left|\begin{array}{cccccccccc} 0 & 2 & & & & & & & & \\ 0 & 0 & 0 & 2 & \ldots & & & & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 2 & 0 & & \\ 2 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & & \\ 0 & 0 & 2 & 0 & \ldots & 0 & 0 & 0 & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 2 & & \end{array}\right| \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} r-1 \\ \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} r. \end{array}$$

Man hat wieder $r(r-1)/2$ Inversionen und folglich ist

$$W_{2r} = (-1)^{r(r-1)/2} \cdot 2^{2r-1}.$$

Für P_n ergibt sich diesmal

$$P_n = P_{2r} = \prod_{\nu=1}^{2r-1} (e^{2\nu\varphi i} + 0) \cdot \prod_{1 \le \nu < \mu \le 2r-1} (e^{2\nu\varphi i} + e^{2\mu\varphi i}).$$

Das erste Produkt ist gleich

$$e^{2r(2r-1)\varphi i} = e^{2r\pi i} = 1,$$

und das zweite ist dasselbe, für das wir oben den Wert $(-1)^{r(r-1)/2}$ fanden. Für die Konstante K_n in (11) ergibt sich also diesmal

$$K_n = K_{2r} = \frac{W_{2r}}{P_{2r}} = 2^{2r-1} = 2^{n-1}.$$

Das Resultat ist daher für gerade und ungerade n das gleiche und es ist $W_n = 2^{n-1} P_n$, was nach Fußnote [6]) auch für den seither ausgeschlossenen Fall $n = 2$ gilt. Somit gewinnen wir die folgende Präzisierung eines Satzes von PÓLYA:

Satz 2. *Sind $s_\nu = x_1^\nu + \cdots + x_n^\nu$ die Potenzsummen von n Variablen, so sind $s_1, s_3, s_5, \ldots, s_{2n-1}$ voneinander unabhängig und jede symmetrische Form ist eine rationale, im allgemeinen gebrochene Funktion von diesen. Insbesondere gilt die Formel*

$$\prod_{\nu<\mu}(x_\nu + x_\mu) = \frac{1}{2^{n-1}} \begin{vmatrix} w_1 & 2 & & & & \\ w_3 & w_2 & w_1 & 2 & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & w_{2n-5} & \cdots & \cdots & w_{n-1} \end{vmatrix} = \frac{W_n}{2^{n-1}}$$

$$(W_1 = 1,\ W_2 = w_1),$$

während die symmetrischen Grundfunktionen $a_1, a_2, \ldots, a_n$ sich durch Auflösung der ersten n Gleichungen (6) ergeben, wobei die Nennerdeterminante gleich $2W_n$ ist. Die w_ν sind dabei Polynome von $s_1, s_3, s_5, \ldots, s_{2n-1}$, die sich rekurrent aus dem Formelsystem (5) berechnen.

Herr PÓLYA hat nicht die Determinante W_n betrachtet (a. a. O.), sondern diese:

$$R_n = \begin{vmatrix} w_1 & w_2 & \cdots & w_n \\ w_2 & w_3 & \cdots & w_{n+1} \\ & & \cdots & \\ w_n & w_{n+1} & \cdots & w_{2n-1} \end{vmatrix}.$$

Er konnte zeigen: Erstens, daß R_n als Polynom von $s_1, s_3, \ldots, s_{2n-1}$ in zwei Faktoren zerfällt, deren einer von s_{2n-1} frei und gleich P_n und deren anderer gleich $(-1)^{n(n+1)/2} \cdot 2^n a_n P_n$ ist, zweitens, daß auch die Produkte $a_\nu P_n$ Polynome von $s_1, s_3, \ldots, s_{2n-1}$ sind; doch hat er für P_n keine formelmäßige Darstellung gegeben (außer für $n = 1, 2, 3, 4$). Vorläufer von PÓLYA ist LAGUERRE[7]).

§ 3. Zusätzliche Bemerkungen.

Bemerkung I. W_n ist als Polynom der s_ν irreduzibel (PÓLYA a. a. O.). Im Fall der Reduzibilität würde nämlich, indem man $n+2$ an Stelle von n schreibt[8]) und dann die Variablen x_{n+1}, x_{n+2} zu 0 spezialisiert, sich auch das Polynom (7) als reduzibel erweisen. Dann müßte einer seiner Faktoren als Polynom der x_ν verschwinden, und das wäre eine Abhängigkeit zwischen $s_1, s_3, \ldots, s_{2n-1}$ von einer Höhe kleiner als H, die es aber nach Satz 1 nicht gibt.

Bemerkung II. Wenn man $a_1, a_2, \ldots, a_n$ durch $s_1, s_2, \ldots, s_n$ ausdrückt, so kommen bei a_ν nur die ersten ν vor: $s_1, \ldots, s_\nu$. Man könnte denken, daß bei der Darstellung durch $s_1, s_3, \ldots, s_{2n-1}$ vielleicht etwas Ähnliches gilt. Das trifft aber nicht zu. Bei a_1 allerdings ist es so, daß nur s_1 vorkommt; es ist ja $a_1 = -s_1$. Bei der Auflösung der ersten n Gleichungen (6) nach a_1 kommt

[7]) LAGUERRE, E.: Sur un problème d'algèbre. Bull. Soc. Math. France **5**, 26—30 (1877) = Oeuvres **I**, 119—122.

[8]) Das darf man, da für $n = 1$ und $n = 2$ die Sache trivial ist: $W_1 = 1$ ist konstant, $W_2 = w_1 = 2s_1$ ist linear.

eben auch im Zähler die Determinante W_n als Faktor vor, der sich dann gegen den Nenner weghebt. Aber schon bei a_2 ist das nicht so; hier kann kein etwaiger Faktor der Zählerdeterminante gegen einen Faktor des Nenners W_n weggekürzt werden. Denn wegen der Irreduzibilität von W_n müßte W_n selbst dieser Faktor sein, und man erhielte a_2 als Polynom von $s_1, s_3, \ldots, s_{2n-1}$; aber das müßte vom Gewicht 2 sein, könnte also nur die Gestalt $C s_1^2$ haben und wäre nicht gleich a_2.

Bemerkung III. Durch die Potenzsummen $s_1, s_2, \ldots, s_n$ sind die Grundfunktionen eindeutig bestimmt, gleichgültig, ob die x_ν unabhängige Variable sind oder irgendwie spezialisiert. Dabei dürfen für $s_1, s_2, \ldots, s_n$ willkürliche Werte vorgegeben sein, es gibt immer zugehörige $a_1, a_2, \ldots, a_n$, und zwar eindeutig. Durch $s_1, s_3, \ldots, s_{2n-1}$ dagegen sind die a_ν nur dann eindeutig bestimmt, wenn die Determinante W_n von 0 verschieden ist. Wenn dagegen die $s_1, s_3, \ldots, s_{2n-1}$ so vorgegeben sind, daß $W_n = 0$ ist, so gibt es zugehörige Systeme $a_1, \ldots, a_n$ dann und nur dann, und zwar unendlich viele, wenn auch alle Zählerdeterminanten, also schließlich alle Determinanten der Matrix

$$\begin{Vmatrix} w_1 & 2 & & & & & \\ w_3 & w_2 & w_1 & 2 & & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-1} & w_{2n-2} & w_{2n-3} & \cdots & \cdots & \cdots & w_{n-1} \end{Vmatrix}$$

verschwinden. Zum Beispiel für $n = 3$ ist

$$W_3 = \begin{Vmatrix} w_1 & 2 \\ w_3 & w_2 \end{Vmatrix} = \tfrac{1}{3}(s_1^3 - s_3).$$

Wird nun etwa $s_1 = c$, $s_3 = c^3$ vorgegeben, so daß W_3 verschwindet, so ist

$$w_1 = 2c,\ w_2 = 2c^2,\ w_3 = 2c^3,\ w_4 = 2c^4,\ w_5 = 2(4c^5 + s_5)/5,$$

und die obige Matrix ist folgende:

$$\begin{Vmatrix} 2c & & 2 & 0 & 0 \\ 2c^3 & & 2c^2 & 2c & 2 \\ 2(4c^5 + s_5)/5 & 2c^4 & 2c^3 & 2c^2 & \end{Vmatrix}.$$

Ihre Determinanten verschwinden sämtlich nur für $s_5 = c^5$. Gibt man also für s_5 einen von c^5 verschiedenen Wert vor, so gibt es keine zugehörigen a_ν, also kein Polynom $F(t)$; wenn aber $s_5 = c^5$ vorgegeben wird, so gibt es unendlich viele; man findet leicht, daß es die Polynome

$$F(t) = (1 - ct)(1 - dt^2)$$

sind mit beliebigem d.

München 27, Friedrich-Herschel-Str. 11.

(Eingegangen am 13. November 1954.)

ROSENTHAL, A.
Math. Zeitschr. Bd. 63, S. 31—38 (1955)

On the continuity of functions of several variables.

In memory of ISSAI SCHUR.

By
ARTHUR ROSENTHAL.

1. Introduction. Let us first consider functions $f(x, y)$ of two variables. It is a well known fact that a function $f(x, y)$ can be continuous along every straight line through the point $p_0 \equiv (x_0, y_0)$ without being continuous at p_0 as a function of (x, y)[1]). Moreover, many years ago, H. LEBESGUE[2]) called attention to the fact that $f(x, y)$ can be discontinuous at p_0 even if $f(x, y)$ is continuous along every analytic curve through p_0[3]). Hence one may ask the following questions: (a) By which more general class of curves can the straight lines or the analytic curves be replaced, i.e., on which class of curves through p_0 may continuity of f be assumed without implying the continuity of $f(x, y)$ at p_0 as a function of (x, y)? (b) On which class of curves through p_0 is the continuity of f sufficient to guarantee the continuity of $f(x, y)$ as a function of (x, y)? The discussion of these questions will give the following result:

Theorem 1. *If the single-valued function f is continuous at p_0 along every convex curve through p_0 which is (at least) once differentiable*[4]), *then f is also continuous at p_0 as a function of (x, y). Yet f can be continuous along every curve through p_0 which is (at least) twice differentiable without being continuous at p_0 as a function of (x, y).*

---

[1]) Probably this was first observed by J. THOMAE, Abriß einer Theorie der complexen Funktionen, 2nd ed., Halle 1873, p. 15; see Encykl. d. Math. Wiss. II A 1 (A. PRINGSHEIM), footnote 254. The following simple example was given in A. GENOCCHI-G. PEANO, Calcolo differentiale e principii di Calcolo integrale, Torino 1884, No. 123:

$$f(x, y) = \begin{cases} \dfrac{x y^2}{x^2 + y^4} & \text{except for } x = y = 0, \\ 0 & \text{for } x = y = 0. \end{cases}$$

[2]) LEBESGUE, H.: J. de Math. (6) 1, 199—200 (1905). Analogous observations concerning maxima and minima were previously made by SCHEEFFER, L.: Math. Ann. 35, 545—551 (1890).

[3]) An example of this kind is the following:

$$f(x, y) = \begin{cases} 1 \text{ on the curve } y = e^{-\frac{1}{x^2}} & \text{except for } x = 0, \\ 0 \text{ otherwise.} \end{cases}$$

A function $f(x, y)$ of this type which, moreover, is continuous everywhere except for $x = y = 0$ can be defined analogously to our formula (2).

[4]) A straight line segment is considered to be a particular case of a convex curve. Furthermore, a differentiable convex arc is also continuously differentiable [see, e.g., O. HAUPT-G. AUMANN, Differential- und Integralrechnung, vol. 2, p. 61, Berlin 1938].

The properties of curves, applied here, are to be understood in the following manner: Let the curve C have a tangent at p_0. Take this tangent-line as ξ-axis with p_0 as origin. With respect to this (ξ, η)-coordinate system and in a sufficiently small neighborhood N of p_0, the curve C is to be represented by a single-valued function $\eta = \varphi(\xi)$ which, in the first case [see[4])], has a continuous derivative everywhere in N and, in the second case, has, moreover, a second derivative, at least at p_0.

The proof of Theorem 1 is obtained by means of the following general

Theorem 2. *Let M be a bounded infinite set of points in the plane. Then M contains an infinite subset M^* through which a convex curve C can be laid which is (at least) once differentiable. But there does not always exist an infinite subset of M through which any (at least) twice differentiable curve can be laid.*

In the expression "(at least) once or twice differentiable" the phrase "(at least)" will be omitted in the following.

In § 4 many other applications of this Theorem 2 will be given. In § 5 and 6 the generalizations to functions f of more than two variables will be discussed. All results are quite analogous to the 2-dimensional case; but the differentiable convex arcs (in the plane) will be replaced by differentiable "primitive" arcs (in spaces of more dimensions). Their definition will be given in § 5.

2. Proofs (first part). We now prove the first part of Theorem 2. Let M be a bounded infinite set of points in the plane and let A be a denumerably infinite subset of M with only one point p_0 of accumulation and such that p_0 does not belong to A. We join p_0 with all the points of A by straight rays and consider their limits. Let g be a ray issuing from p_0, which occurs as such a limit. If there are infinitely many points of A on g, then the set of these points can be taken as M^* with g representing C. Otherwise there is a subsequence of points $\{a_\nu\}$ of A, not lying on g, such that the rays $\overrightarrow{p_0 a_\nu}$ converge to g. Then we choose p_0 as the origin of a rectangular coordinate-system with g as the positive x-axis. Let (x_ν, y_ν) be the coordinates of a_ν. By $(a\,b\,x)$ we designate the angle $(\leq \pi)$ between the directed straight line $\overrightarrow{b\,a}$ and the positive x-axis. Then we can choose a subsequence $\{a_{\nu_k}\}$ of $\{a_\nu\}$ such that $x_{\nu_k} > x_{\nu_{k+1}} > 0$ and such that the angles $(a_{\nu_k}\, p_0 x)$ are acute and the sequence of these angles is strictly decreasing. The angles $(a_{\nu_k} a_{\nu_{k+m}} x)$ (for $m = 1, 2, \ldots$) are greater than the angle $(a_{\nu_k} p_0 x)$, but converge to it as $m \to \infty$. Hence from $\{a_{\nu_k}\}$ we can select a new subsequence, which we call $\{b_n\}$, such that the angles $(b_n b_{n+1} x)$ strictly decrease to 0 as n increases. Then the points b_n are the vertices of a generalized convex polygon with denumerably many sides converging to p_0 and with the positive x-axis as half-tangent at p_0. At every vertex b_n (with $n \geq 2$) we draw the bisector g_n of the supplement of the interior angle $b_{n-1} b_n b_{n+1}$. Let the intersection of g_n and g_{n+1} be designated by c_n. The acute angle γ_n determined by g_n and the x-axis is the arithmetic mean of the angles $(b_{n-1} b_n x)$ and $(b_n b_{n+1} x)$, and hence the γ_n

decrease strictly to 0 as $n \to \infty$. Therefore the generalized polygon $\mathfrak{P}$ with the vertices c_n $(n = 2, 3, \ldots)$ is also convex. Its sides $c_n c_{n+1}$ (containing b_{n+1}) converge to p_0, again with the positive x-axis as half-tangent of $\mathfrak{P}$ at p_0. If we round off every vertex c_n of $\mathfrak{P}$ by a sufficiently small circular arc which touches the sides of the vertex c_n, then we obtain a differentiable convex arc C with p_0 as endpoint and the positive x-axis as half-tangent at p_0. Joining C with its reflection in the y-axis gives a differentiable convex arc with the x-axis as tangent at p_0. This proves the first part of theorem 2.

The first part of theorem 1 now follows immediately. In fact, let the values of f be continuous at $p_0 \equiv (x_0, y_0)$ on every differentiable convex curve through p_0. Then if $f(x, y)$ were not continuous at p_0 as a function of (x, y), there would be a sequence of points $\{a_\nu\}$ converging to p_0 for which $\lim_{\nu \to \infty} f(x_\nu, y_\nu)$ would exist, but would be different from $f(x_0, y_0)$. Then, according to the above discussion, there would be a subsequence $\{a_{\nu_n}\}$ through which one could lay a differentiable convex curve C. Hence, by assumption, we would have $\lim_{n \to \infty} f(x_{\nu_n}, y_{\nu_n}) = f(x_0, y_0)$, in contradiction to the selection of the $\{a_\nu\}$.

3. Proofs (second part). We shall now show that a bounded infinite set M does not always contain an infinite subset through which a twice differentiable curve can be laid. For this purpose we consider a continuous curve $y = f(x)$ passing through 0 with $f'(0) = 0$ and with the improper second derivative $f''(0) = +\infty$; e.g., for $y = f(x)$ we take the semicubical parabola $y = a \cdot |x|^{\frac{3}{2}}$ (with $a > 0$). On the branch of this curve for which $x > 0$ let $\{a_\nu\}$ be a sequence of points converging to the origin and let this sequence be our set M. Then no twice differentiable curve C can be laid through any infinite subset of M. For let $y = \varphi(x)$ with $\varphi(0) = 0$ and $\varphi'(0) = 0$ represent such a curve C. Then since $\lim_{h \to 0^+} \frac{\varphi'(h) - \varphi'(0)}{h} = L$ (finite), for every $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that for $0 < h < \delta_1$ we have $|L| - \varepsilon < \left|\frac{\varphi'(h)}{h}\right| < |L| + \varepsilon$. On the other hand, to $Q > |L| + \varepsilon$ there exists a $\delta_2 > 0$ such that for $0 < h < \delta_2$ we have $\frac{f'(h)}{h} > Q$ (i.e., in our case, $\frac{3}{2} a \cdot h^{-\frac{1}{2}} > Q$). Thus if $\delta = \min(\delta_1, \delta_2)$, for $0 < h < \delta$ we have $|\varphi'(h)| < f'(h)$, and hence by integration $|\varphi(h)| < f(h)$ for $0 < h < \delta$ (where if f is given, δ depends on φ). Therefore no twice differentiable curve $y = \varphi(x)$ can go through infinitely many points of M. This proves the second part of theorem 2.

The second part of theorem 1 is then an immediate consequence, since one can now easily give examples of functions $f(x, y)$ which are continuous along every twice differentiable curve through (0, 0) without being continuous at the origin as functions of (x, y). For instance set

$$(1) \qquad f(x, y) = \begin{cases} 1 \text{ on the curve } y = |x|^{\frac{3}{2}} \text{ except for } x = 0, \\ 0 \text{ otherwise } [(0, 0) \text{ included}]. \end{cases}$$

In order to obtain such a function which is continuous everywhere except for (0, 0), one can proceed in the following manner. Let C_1, C_2, C_3 be three

curves with the equations $y=a|x|^{\frac{1}{2}}$, $y=b|x|^{\frac{1}{2}}$, $y=c|x|^{\frac{1}{2}}$, respectively, where $b>c>a>0$. In the "strip" S bounded by C_1 and C_2, let there be a "mountain" above the xy-plane whose crest lies above C_3 with a constant altitude h, except at the origin, and which slopes down to the curves C_1 and C_2. We now define the single-valued function $z=f(x, y)$ correspondingly: Inside S it is represented by the "mountain", while it vanishes on C_1 and C_2 (the origin included) and everywhere outside of S. A very simple definite example of such a function $f(x, y)$ is the following:

$$(2)\quad f(x,y)=\begin{cases}|(y|x|^{-\frac{1}{2}}-a)\,(y|x|^{-\frac{1}{2}}-b)|-(y|x|^{-\frac{1}{2}}-a)\,(y|x|^{-\frac{1}{2}}-b) & \text{for } x\neq 0,\\ 0 & \text{for } x=0\end{cases}$$

with $b>a>0$. Here the curve C_3 is determined by $y=\frac{1}{2}(a+b)|x|^{\frac{1}{2}}$ and on C_3, except for $x=0$, f has the constant value $h=\frac{1}{2}(b-a)^2$.

These examples accomplish what we desire. In fact, every twice differentiable curve C passing through the origin 0 lies outside of S in a sufficiently small deleted neighborhood N of 0, and hence f vanishes everywhere on C in N. This is also the case for every such curve C whose tangent at 0 is the x-axis, since by the above such a C (except for 0) lies below C_1 in a sufficiently small neighborhood of 0.

We also note that a function $F(x, y)$ may be discontinuous at the points of a set dense in a region R, although F is continuous on every curve K in R which is twice differentiable at each of its points. Example: Let (x_n, y_n) $(n=1, 2, \ldots)$ be a sequence of all the points in R with both coordinates rational and set

$$(3)\qquad F(x,y)=\sum_{n=1}^{\infty}\frac{1}{2^n}f(x-x_n, y-y_n)$$

where $f(x, y)$ denotes the function defined in (2).

4. Other applications. Theorem 2 allows many other applications also.

(a) Let $z=F(x, y)$ represent a continuous surface S and let p_0 be a point on S. It is possible that every curve on S through p_0, such that the projection on the xy-plane is a twice differentiable curve, has a tangent at p_0 and that all these tangents lie in a plane, but that, nevertheless, this plane is not a *tangent plane* of S at p_0. Yet if the same holds for every curve on S through p_0 which is projected on the xy-plane as a differentiable convex curve, then that plane is a tangent plane of S at p_0.

An example for the first part of this statement is given by

$$(4)\qquad F(x,y)=f(x,y)\cdot|x|,$$

where $f(x, y)$ is the function defined by (2).

(b) A continuous function $w=\varphi(z)$, where $z=x+iy$, can have a *derivative* with the same value at z_0 along every twice differentiable curve through z_0, although $\varphi'(z_0)$ does not exist. But $\varphi'(z_0)$ certainly exists if $\varphi(z)$ has a derivative with the same value at z_0 along every differentiable convex curve.

An example for the first part of this statement is obtained from the function $F(x, y)$, defined in (4), by setting $\varphi(x+iy)=F(x, y)$.

(c) If the single-valued function $g(x, y)$ is bounded along every differentiable convex curve through the point p_0, then g is *bounded* in a sufficiently small neighborhood of p_0. Yet g can be bounded along every twice differentiable curve through p_0 without being bounded in any neighborhood of p_0.

An example for the last part of this statement is given by

$$g(x, y)=\begin{cases} f(x, y)\cdot\frac{1}{|x|} & \text{for } x\neq 0, \\ 0 & \text{for } x=0, \end{cases} \tag{5}$$

where $f(x, y)$ again designates the function defined by (2).

(d) Let the single-valued function $f(x, y)$ be defined in a deleted neighborhood of $p_0\equiv(x_0, y_0)$. Then

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) \tag{6}$$

certainly exists if f has the *same* limit at p_0 along every differentiable convex curve through p_0. But even if f has the same limit at p_0 along every twice differentiable curve through p_0, then the limit (6) need not exist.

Here it is essential that f has always the *same* limit at p_0. For even the existence of a limit of f at p_0 (but not necessarily always the same limit) along each differentiable curve through p_0 does not guarantee the existence of (6). Example: Let S be the part of the screw surface

$$x=r\cos\vartheta, \qquad y=r\sin\vartheta, \qquad z=\vartheta, \qquad 0\leq\vartheta\leq\frac{\pi}{2}$$

and let $z=f(x, y)$ represent the surface obtained by continuing S by its reflections in the xz- and yz-planes.

(e) The single-valued function $f(x, y)$ has a *maximum* at p_0 if f has a maximum at p_0 along every differentiable convex curve through p_0. But f can have a maximum at p_0 along every twice differentiable curve through p_0 without having a maximum at p_0 as a function of (x, y). Example:

$$f(x, y)=(y-a|x|^{\frac{1}{2}})\cdot(b|x|^{\frac{1}{2}}-y) \quad \text{with } b>a>0. \tag{7}$$

(f) Let $\{f_\nu(x, y)\}$ be a sequence of single-valued functions defined in a neighborhood of p_0. Then the *uniform convergence* of $\{f_\nu(x, y)\}$ at p_0 is implied by the uniform convergence of $\{f_\nu\}$ at p_0 along every differentiable convex curve through p_0; but it is not necessarily implied by the uniform convergence of $\{f_\nu\}$ at p_0 along every twice differentiable curve through p_0.

Proof of the first part: Suppose $f_\nu(x, y)$ is not uniformly convergent at p_0. Then there exists an $\varepsilon_0>0$, a sequence of points $p_n\equiv(x_n, y_n)$ converging to p_0, and, for every n, two indices ν_n and ν_n^* with $\nu_n^*>\nu_n\geq n$, such that

$$|f_{\nu_n^*}(x_n, y_n)-f_{\nu_n}(x_n, y_n)|\geq\varepsilon_0.$$

But then there is a differentiable convex arc C through a subsequence of $\{p_n\}$ along which $\{f_\nu\}$ is not uniformly convergent at p_0.

Example for the second part: Set

$$(8)\qquad \psi_\nu(\varrho)=\begin{cases} 2\nu\varrho & \text{for } 0\le\varrho\le\frac{1}{2\nu},\\ 2(1-\nu\varrho) & \text{for } \frac{1}{2\nu}\le\varrho\le\frac{1}{\nu},\\ 0 & \text{for } \varrho\ge\frac{1}{\nu}\end{cases}$$

and

$$(9)\qquad f_\nu(x,y)=f(x,y)\cdot\psi_\nu\left(\sqrt{x^2+y^2}\right)\qquad(\nu=1,2,\ldots)$$

where $f(x,y)$ is the function defined by (2).

Moreover, we give an example of a sequence $\{F_\nu(x,y)\}$ of continuous functions in a region R, uniformly convergent along every curve in R which is twice differentiable at each of its points, without being uniformly convergent in any subregion of R: Let (x_n, y_n) $(n=1,2,\ldots)$ be a sequence of all the points with both coordinates rational. Then we set

$$(10)\qquad F_\nu(x,y)=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot f_\nu(x-x_n,y-y_n)\qquad(\nu=1,2,\ldots)$$

where the $f_\nu(x,y)$ are the functions defined by (9).

(g) Let $\{f_\nu(x,y)\}$ be a sequence of single-valued functions defined in a bounded region R. If $\{f_\nu\}$ is uniformly bounded along every differentiable convex curve in R, then $\{f_\nu\}$ is *uniformly bounded* in R. But $\{f_\nu\}$ can be uniformly bounded along every curve in R which is twice differentiable at each of its points without being uniformly bounded in R.

Example for the second part:

$$(11)\qquad f_\nu(x,y)=\begin{cases} f(x,y)\cdot(x^2+y^2)^{-\frac{1}{4}}\cdot\psi_\nu\left(\sqrt{x^2+y^2}\right) & \text{except for } x=y=0,\\ & (\nu=1,2,\ldots)\\ 0 & \text{for } x=y=0\end{cases}$$

where $f(x,y)$ and $\psi_\nu(\varrho)$ are the functions defined by (2) and (8).

(h) The situation is quite different for a sequence $\{f_\nu(z)\}$ of regular *analytic* functions in a region R. In order that $\{f_\nu(z)\}$ be *uniformly bounded* in every bounded, closed subregion of R, it is necessary and sufficient that $\{f_\nu(z)\}$ be uniformly bounded *on every closed straight segment of* R. For then $\{f_\nu(z)\}$ is uniformly bounded on every simple closed polygon $\mathfrak{P}$ contained in R, and hence, by the maximum modulus theorem, also in the closed region bounded by $\mathfrak{P}$.

The analogous statement holds also for the *uniform convergence* of $\{f_\nu(z)\}$. This results either from the preceding by means of the Vitali theorem or, quite directly, using the uniform convergence of $\{f_\nu(z)\}$ on $\mathfrak{P}$ and then concluding[5]), by means of the Cauchy integral formula, that $\{f_\nu(z)\}$ converges uniformly also in the region bounded by $\mathfrak{P}$.

[5]) Cf. RUNGE, C.: Acta math. 6, 232—234 (1884). — BIEBERBACH, L.: Lehrbuch der Funktionentheorie, vol. I, p. 166—167, Leipzig u. Berlin 1921 (4th. ed., 1934, p. 168—169).

5. Primitive arcs. Here and in § 6 we shall discuss the generalization to functions of more than two variables. For this purpose we first define (by induction for $n \geq 2$) a *primitive arc*, which in R_n will play the same role as the convex arc in the plane. In R_2 a primitive arc is a convex arc. Suppose the notion of a primitive arc in R_{n-1} has been defined; then we define it in R_n in the following manner: Let C be an arc in R_n through the point p_0 with the tangent t_0 at p_0. Let $\pi_0^{(n-1)}$ be the R_{n-1} through p_0 perpendicular to t_0. Then the orthogonal projection C' of C on $\pi_0^{(n-1)}$ is assumed to be a primitive arc with a tangent t_1 at p_0, and the orthogonal projection C'' of C on the plane (t_0, t_1)[6]) is assumed to be a convex arc. The special case is admitted in which C' shrinks to the point p_0; in this case C is a straight line segment (of t_0). According to this recursive definition there are $n-1$ orthogonal planes through p_0, on which the projections of C are convex arcs (provided that none of them shrinks to p_0).

The primitive arc C is *differentiable* if C has a tangent at each of its points. This is the case for a sufficiently small C if and only if each of the projections on the orthogonal planes referred to above is a differentiable convex arc. In fact, C is then even continuously differentiable (i.e., the tangent varies continuously).

We now state what we mean by a *twice differentiable* arc C in R_n. Let C be a continuously differentiable arc through p_0 with the tangent t_0 at p_0. In a sufficiently small neighborhood N of p_0, the projection C' of C on any plane π through t_0 is a continuously differentiable arc. Taking t_0 as ξ-axis with origin at p_0, let a (ξ, η)-coordinate system be introduced in π; then C' is represented in N by a single-valued, continuously differentiable function $\eta = \varphi(\xi)$. We call C "twice differentiable" if, in each such plane π, at least $\varphi''(0)$ exists.

6. Generalization to n dimensions. Now we generalize Theorem 2 for the n-dimensional space R_n.

Theorem 3. *Let M be a bounded infinite set of points in R_n. Then M contains an infinite subset M^* through which a differentiable primitive arc C can be laid. But there does not always exist an infinite subset of M through which any twice differentiable arc can be laid.*

Proof: Let A be a denumerable subset of M with only one point p_0 of accumulation and let p_0 not belong to A. As at the beginning of § 2 either there is a line g containing infinitely many points of A (then g can be taken as C) or otherwise we can select a subsequence $\{a_\nu^1\}$ of A such that the rays $\overline{p_0 a_\nu^1}$ converge to a limiting ray g_1. Then consider the R_{n-1} perpendicular to g_1 at p_0 and the orthogonal projections $\tilde{a}_\nu^1$ of the a_ν^1 on this R_{n-1}. There is a subsequence $\{a_\nu^2\}$ of $\{\tilde{a}_\nu^1\}$ such that the rays $\overline{p_0 a_\nu^2}$ converge to a limiting ray g_2. (The degenerate case that infinitely many $\tilde{a}_\nu^1$ lie on a ray will be discussed subsequently.) Then consider the R_{n-2} (contained in our R_{n-1}) perpendicular

[6]) This is a "tangential osculating plane" of C at p_0; [cf. O. HAUPT, J. f. Math. **170**, 157 (1934)].

to g_2 at p_0 and the orthogonal projections $\tilde{a}_\nu^2$ of the a_ν^2 on this R_{n-2}. Again there is a subsequence $\{a_\nu^3\}$ of the $\{\tilde{a}_\nu^2\}$ such that the rays $\overline{p_0 a_\nu^3}$ converge to a limiting ray g_3. In this way we continue and finally come to a sequence $\{a_\nu^{n-1}\}$ of projections on a plane R_2 perpendicular to g_{n-2} at p_0, such that the rays $\overline{p_0 a_\nu^{n-1}}$ converge to a limiting ray g_{n-1}. Then let g_n be a ray in this R_2, perpendicular to g_{n-1} at p_0. The subsequence of $\{a_\nu^1\}$ which is projected into $\{a_\nu^{n-1}\}$ is designated by $\{q_\nu\}$.

It is possible that the preceding procedure stops at an earlier moment. There may be a first ray $g_k (2 \leq k \leq n-1)$ which contains infinitely many of the projections $\tilde{a}_\nu^{k-1}$; then we take these points as the subsequence $\{a_\nu^k\}$. The corresponding subsequence of $\{a_\nu^1\}$ (which we may call $\{q_\nu\}$ again) is then already contained in the k-dimensional subspace of R_n which is spanned by $g_1, g_2, \ldots, g_k$. Hence our procedure stops here, and we have to consider only the k-dimensional case.

If our procedure does not stop at an earlier stage, then the rays $g_1, g_2, \ldots, g_n$ determine a rectangular $(x_1, x_2, \ldots, x_n)$-coordinate system in R_n. The projections of $\{q_\nu\}$ on the (x_k, x_{k+1})-plane are denoted by $\{q_\nu^k\}$, $(k=1, 2, \ldots, n-1)$. [Hence $\{q_\nu^{n-1}\}$ is the same as $\{a_\nu^{n-1}\}$.] According to theorem 2, in the (x_1, x_2)-plane we lay a differentiable convex curve C_1 through an infinite subsequence $\{b_\nu^1\}$ of $\{q_\nu^1\}$. Let $\{q_\nu^{2^0}\}$ be the subsequence of $\{q_\nu^2\}$, corresponding to $\{b_\nu^1\}$. Then in the (x_2, x_3)-plane we lay a differentiable convex curve C_2 through an infinite subsequence $\{b_\nu^2\}$ of $\{q_\nu^{2^0}\}$. Let $\{q_\nu^{3^0}\}$ be the subsequence of $\{q_\nu^3\}$, corresponding to $\{b_\nu^2\}$. In this manner we continue until in the (x_{n-1}, x_n)-plane we obtain a differentiable convex curve C_{n-1} through an infinite subsequence $\{b_\nu^{n-1}\}$ of a subsequence $\{q_\nu^{n-1^0}\}$ of $\{q_\nu^{n-1}\}$. Let $\{p_\nu\}$ be an infinite subsequence of $\{q_\nu\}$ whose projections on each (x_k, x_{k+1})-plane are included in the $\{b_\nu^k\}$ $(k=1, 2, \ldots, n-1)$. Then by $C_1, C_2, \ldots, C_{n-1}$ a differentiable primitive arc C is determined going through the points $\{p_\nu\}$. In fact, C is the intersection of $n-1$ projecting cylinders.

The second part of theorem 3 is shown by the following example: Let C be the curve in R_n, which is determined by the curves $x_{k+1} = |x_k|^{\frac{1}{2}}$ in the (x_k, x_{k+1})-planes $(k=1, 2, \ldots, n-1)$. Then let M be a sequence of points on C, which converges to the origin.

An immediate consequence of Theorem 3 is the following

Theorem 4. *One can generalize Theorem 1 to single-valued functions $f(x_1, x_2, \ldots, x_n)$ if there one replaces "convex curve" by "primitive curve".*

One also obtains easily other applications of Theorem 3, analogous to those discussed in § 4.

Dept. of Math., Purdue University, Lafayette, Indiana (U.S.A.).

(Eingegangen am 12. Dezember 1954.)

KNOPP, K.
Math. Zeitschr. Bd. 63, S. 39—52 (1955)

Nörlund-Verfahren für Funktionen*).

(Zweite Mitteilung.)

Dem Andenken an ISSAI SCHUR gewidmet.

Von

KONRAD KNOPP.

1. Einleitung.

Die funktionalen Analoga zu den bekannten NÖRLUNDschen Verfahren zur Limitierung von Zahlenfolgen sind in einer vorangehenden gemeinsamen Arbeit des Verfassers mit B. VANDERBURG [1][1]) nach verschiedenen Richtungen untersucht worden. Dort wurde die Klasse S aller (komplexwertigen) Funktionen $s(t)$ eingeführt, die in $t>0$ erklärt und in jedem Intervall $(0, c)$, $c>0$, meßbar und beschränkt sind, sowie die Klasse $\mathcal{P}$ aller (reellwertigen) Gewichtsfunktionen $p(t)$, die in $t>0$ erklärt, in jedem Intervall (a, b), $0<a<b<\infty$, meßbar und beschränkt sind und für die die Integrale

$$P(x)=\int_0^x p(t)\,dt \tag{1.1}$$

bei 0 absolut konvergieren und für alle $x>0$ positiv sind. Wenn $s\in S$ und $p\in\mathcal{P}$, so heißt die für alle $x>0$ erklärte Funktion

$$\sigma(x)=\frac{1}{P(x)}\int_0^x p(x-t)\,s(t)\,dt \tag{1.2}$$

das N_p-*Mittel* von $s(t)$. Strebt $\sigma(x)\to s$ für $x\to\infty$, so heißt $s(t)$ N_p-*limitierbar* zum Werte s für $t\to\infty$. Dies Verfahren heißt *konvergenztreu*, wenn die Konvergenz von $s(t)$ diejenige von $\sigma(x)$ zur Folge hat[2]), es heißt *multiplikativ*, wenn überdies $\lim\sigma(x)=a\cdot\lim s(t)$ ist, und *regulär*, wenn der Faktor $a=1$ ist. Ist $p(t)\geq 0$, so sagen wir kurz, N_p sei positiv.

Es werde ferner die Teilklasse $\mathcal{V}\subset S$ aller Funktionen $s(t)\in S$ eingeführt, die in $(0,\infty)$ von beschränkter Gesamtschwankung sind. Wenn (1.2) jedes $s(t)\in\mathcal{V}$ in ein $\sigma(x)\in\mathcal{V}$ transformiert, so werden die eben erklärten Verfahren *absolut* konvergenztreu, multiplikativ bzw. regulär genannt.

---

*) This research was supported by the United States Air Force through the Office of Scientific Research, of the Air Research and Development Command, under Contract Number AF 18(600)—691.

[1]) Auf diese Arbeit wird im folgenden mit „I" verwiesen.

[2]) Die Grenzwerte sind, wenn nichts anderes ausdrücklich gesagt wird, stets für $t\to\infty$ bzw. $x\to\infty$ genommen.

Liegt ein Integral $\int_0^\infty a(u)\,du$ vor mit $a \in S$ und ist $\int_0^t a(u)\,du = s(t)$ N_p-limitierbar zum Werte s, so heißt das vorgelegte Integral *N_p-summierbar* zum Werte s, insbesondere *absolut* N_p- oder $|N_p|$-summierbar zum Werte s, wenn noch $\sigma(x) \in \mathcal{V}$ gilt.

In § 2 werden zwei Sätze über allgemeine Integraltransformationen bewiesen und in § 3 zwei Sätze über absolute Inklusion (Erklärung s. § 3), die selbständiges Interesse haben. Sie setzen uns instand, in § 4 mehrere Multiplikationssätze zu beweisen, d.h. Sätze, die unter verschiedenen Voraussetzungen auf folgende Frage antworten: Wenn die Integrale $\int_0^\infty a(t)\,dt$ und $\int_0^\infty b(t)\,dt$ je durch ein bestimmtes unserer Verfahren summierbar sind und wenn

$$a * b = \int_0^t a(t-u)\,b(u)\,du = c(t) \tag{1.3}$$

gesetzt wird[3]), — was läßt sich dann über die Summierbarkeit von $\int_0^\infty c(t)\,dt$ sagen? Es werden drei Antworten auf diese Frage gegeben.

Für den Fall unendlicher Reihen $\sum a_\nu$, $\sum b_\nu$, $\sum c_\nu$, bei denen $c_\nu = \sum_{\varrho=0}^{\nu} a_{\nu-\varrho} b_\varrho$ ist, wurden die Antworten von FL. M. MEARS [2], [3], [4], [5] gegeben. Die folgenden Beweise lehnen sich an eine vereinfachte, noch unveröffentlichte Darstellung von G. G. LORENTZ und dem Verfasser an.

2. Zwei Sätze über Integraltransformationen.

Wir brauchen zwei Sätze über Integraltransformationen, die auch selbständiges Interesse haben.

Satz 1. *Die Integraltransformation*[4])

$$\sigma(x) = \int_0^\infty a(x,t)\,s(t)\,dt, \qquad s(t) \in \mathcal{V}, \tag{2.1}$$

ist absolut konvergenztreu, bildet also $\mathcal{V}$ in $\mathcal{V}$ ab, wenn die beiden folgenden Bedingungen erfüllt sind:

$$\int_0^\infty a(x,t)\,dt = A(x) \quad \textit{existiert für jedes} \quad x \geq 0, \tag{2.2}$$

und wird

$$\int_t^\infty a(x,u)\,du = \varphi(x,t), \qquad (x,t \geq 0), \tag{2.3}$$

---

[3]) Vgl. I, § 3.

[4]) Das Integral rechts in (2.1) und in analogen Fällen ist als $\lim \int_0^T$ für $T \to \infty$ zu verstehen.

gesetzt, so ist die Gesamtschwankung von $\varphi(x,t)$ über $(0,\infty)$ bezüglich x gleichmäßig beschränkt in t, d.h. es gibt ein festes $M>0$, so daß

$$\int_0^\infty |d_x\varphi(x,t)| \leq M \quad \text{für alle} \quad t \geq 0 \text{ [5]}. \tag{2.4}$$

Beweis. Da $s(t) \in \mathcal{V}$, ergibt sich durch partielle Integration

$$\sigma(x) = \int_0^\infty \varphi(x,t)\, ds(t) + s(0)\, A(x). \tag{2.5}$$

Auf Grund des Lemmas in I, § 6 und weil $A(x) = \varphi(x,0)$ ist, ergibt sich nun sofort

$$\int_0^\infty |d\sigma(x)| \leq M \left\{ |s(0)| + \int_0^\infty |ds(t)| \right\}. \tag{2.6}$$

Daher ist $\sigma(x) \in \mathcal{V}$.

Bemerkungen. Aus den Voraussetzungen von Satz 1 folgt:

1. Es existiert

$$\lim A(x) = a. \tag{2.7}$$

Denn die Gesamtschwankung von $A(x) = \varphi(x,0)$ über $(0,\infty)$ ist $\leq M$.

2. Für jedes feste $t \geq 0$ existiert

$$\lim_{x\to\infty} \int_0^t a(x,u)\, du = a^*(t) \text{ [6]}. \tag{2.8}$$

Denn da die Gesamtschwankung von $\varphi(x,t)$ bezüglich x endlich ist, existiert zunächst $\lim \varphi(x,t)$ für $x \to \infty$ bei jedem festen $t \geq 0$. Daraus folgt aber die Behauptung, weil die linke Seite von (2.8) mit $\lim \left(\varphi(x,0) - \varphi(x,t)\right)$ für $x \to \infty$ äquivalent ist.

3. Wenn überdies für jedes $T>0$ der Grenzwert (2.8) in $0 \leq t \leq T$ gleichmäßig existiert und wenn $|\varphi(0,t)| \leq M$ [7] ist für alle $t \geq 0$, dann ist

$$\lim_{x\to\infty} \sigma(x) = a \cdot s - \int_0^\infty a^*(t)\, ds(t). \tag{2.9}$$

Nach (2.4) ist nämlich für jedes $X>0$ und alle $t \geq 0$

$$\int_0^X |d_x\varphi(x,t)| \leq M,$$

folglich

$$|\varphi(X,t) - \varphi(0,t)| \leq M,$$

---

[5]) Die beiden Bedingungen (2.2) und (2.4) sind auch notwendig dafür, daß (2.1) absolut konvergenztreu ist. Einen Beweis dafür hat K. ZELLER für den Verfasser ausgearbeitet. Doch da im folgenden von der Notwendigkeit der Bedingungen kein Gebrauch gemacht wird, unterdrücken wir den nicht ganz einfachen Beweis. — Vgl. hierzu auch: G. SUNOUCHI und T. TSUCHIKURA [6].

[6]) Im Fall von Matrixtransformationen entspricht dies der Existenz der Spaltenlimites.

[7]) O. B. d. A. darf hier das gleiche M wie in (2.4) genommen werden.

also

$$|\varphi(x,t)| \leq 2M \quad \text{für alle} \quad x,t \geq 0$$

und

$$\left|\int_0^t a(x,u)\,du\right| = |\varphi(x,0) - \varphi(x,t)| \leq 4M.$$

Daher sind alle Voraussetzungen des nachfolgenden Satzes 2 erfüllt [mit $4M$ statt M in (2.10)], so daß die jetzige Behauptung (2.9) in der Behauptung (2.13) dieses Satzes 2 enthalten ist, den wir sogleich beweisen.

4. Wenn $a^*(t) \equiv 0$ ist, so ist (2.1) absolut multiplikativ mit dem Faktor $a = \lim A(x)$ und absolut regulär, wenn noch $a = 1$ ist.

Satz 2. *Die Funktion $\sigma(x)$ in (2.1) konvergiert für jedes $s(t) \in \mathcal{V}$, falls mit einem von x und t unabhängigen M*

$$\left|\int_0^t a(x,u)\,du\right| \leq M \tag{2.10}$$

ist für alle $x,t \geq 0$,

$$A(x) = \int_0^\infty a(x,t)\,dt$$

für alle $x \geq 0$ existiert und

$$A(x) \to a \quad \text{strebt für} \quad x \to \infty \tag{2.11}$$

und wenn für jedes $T > 0$ der Grenzwert

$$\lim_{x\to\infty} \int_0^t a(x,u)\,du = a^*(t) \tag{2.12}$$

gleichmäßig in $0 \leq t \leq T$ existiert. Sind diese Bedingungen erfüllt, so ist überdies

$$\lim_{x\to\infty} \sigma(x) = a \cdot s - \int_0^\infty a^*(t)\,ds(t) = \sigma. \tag{2.13}$$

Beweis. Mit der Bezeichnung (2.3) ist zunächst nach (2.10) und (2.11)

$$|\varphi(x,t)| \leq 2M \quad \text{für alle} \quad x,t \geq 0$$

und es strebt für $x \to \infty$

$$\varphi(x,t) \to a - a^*(t) \tag{2.14}$$

gleichmäßig in $0 \leq t \leq T$, so daß auch

$$|a - a^*(t)| \leq 2M \quad \text{ist für alle} \quad t > 0. \tag{2.15}$$

Wird also bei gegebenem $\varepsilon > 0$ das T so gewählt, daß

$$\int_T^\infty |ds(t)| < \frac{\varepsilon}{8M}$$

ist und wird

$$\begin{aligned}\sigma(x) - \sigma &= \int_0^T + \int_T^\infty \{\varphi(x,t) - a + a^*(t)\}\,ds(t) + \{A(x) - a\}\,s(0)\\ &= A + B + C\end{aligned}$$

gesetzt, so ist zunächst $|B|<\varepsilon/2$. Ferner ist

$$A=\int_0^T\left\{A(x)-\int_0^t a(x,u)\,du-a+a^*(t)\right\}ds(t),$$

also

$$|A|\leq|A(x)-a|\int_0^\infty|ds(t)|+\int_0^T\left|\int_0^t a(x,u)\,du-a^*(t)\right|\cdot|ds(t)|<\varepsilon/4$$

für alle hinlänglich großen x, weil rechterhand beide Summanden $\to 0$ streben für $x\to\infty$, der erste nach (2.11), der zweite nach (2.12). Da schließlich auch $|C|<\varepsilon/4$ ist für alle hinreichend großen x, so ist $|A+B+C|<\varepsilon$ für $x>x_0(\varepsilon)$. Damit ist die Behauptung (2.13) bewiesen.

3. Ein Satz über absolute Inklusion.

Die Gewichtsfunktionen p, q und $p*q=r$ mögen zu $\mathcal{P}$ gehören und es werde wie üblich

$$\int_0^x p(t)\,dt=P(x),\qquad\int_0^x r(t)\,dt=R(x)$$

gesetzt. Dann gilt der

Satz 3. *Jede $|N_p|$-limitierbare Funktion $s(t)\in S$ ist auch $|N_r|$-limitierbar, falls für ein von t unabhängiges M und alle $t\geq 0$ die Gesamtschwankung*

$$V_{x-t}^{\infty}\frac{1}{R(x)}\int_0^t P(u)\,q(x-u)\,du\leq M \tag{3.1}$$

ist. Und wenn überdies für jedes feste $t>0$

$$\frac{1}{R(x)}\int_0^t|q(x-u)|\,du\to 0 \tag{3.2}$$

strebt für $x\to\infty$, so ist auch N_p-lim $s(t)=N_r$-lim $s(t)$, d.h. das Verfahren $|N_p|$ ist in dem Verfahren $|N_r|$ enthalten (in Zeichen: $|N_p|\subseteq|N_r|$).

Beweis. Wir setzen

$$N_p\big(x,s(t)\big)=\sigma(x),\qquad N_r\big(x,s(t)\big)=\tau(x).$$

Dann ist (vgl. I, § 3) $p*s=P\cdot\sigma$ und also $r*s=q*(P\cdot\sigma)$, d.h.

$$\tau(x)=\int_0^\infty a(x,t)\,\sigma(t)\,dt\qquad\text{mit}\qquad a(x,t)=\frac{q(x-t)\,P(t)}{R(x)}. \tag{3.3}$$

Für die erste Behauptung des Satzes genügt es nun zu zeigen, daß diese Transformation (3.3) von $\sigma(t)$ in $\tau(x)$ absolut konvergenztreu ist oder daß sie die Bedingungen (2.2) und (2.4) erfüllt. Da aber $\int_0^\infty a(x,t)\,dt=A(x)\equiv 1$

ist, so ist (2.2) gewiß erfüllt. Und im jetzigen Falle ist ersichtlich

$$\varphi(x,t) = 1 - \frac{1}{R(x)} \int_0^t P(u)\, q(x-u)\, du,$$

so daß (2.4) mit (3.1) äquivalent ist.

Für die zweite Behauptung des Satzes ziehen wir die Bemerkungen 2 und 3 zu Satz 1 heran. Es ist jetzt

$$a^*(t) = \lim_{x\to\infty} \int_0^t a(x,u)\, du = \lim_{x\to\infty} \frac{1}{R(x)} \int_0^t P(u)\, q(x-u)\, du,$$

und zwar für jedes $T>0$ gleichmäßig in $0 \leq t \leq T$, weil

$$\int_0^t P(u)\, q(x-u)\, du \leq \max_{0\leq u\leq T} |P(u)| \cdot \int_0^T |q(x-u)|\, du = o\big(R(x)\big)$$

ist, letzteres nach (3.2). Da schließlich $a=1$ ist, so ist (3.3) absolut regulär.

Der folgende Satz gibt etwas einfachere Bedingungen für dieselbe Inklusion:

Satz 4. *Für die in Satz 3 behauptete Inklusion ist es hinreichend, wenn*

(3.4) $q(t)$ *nicht negativ*

ist und für jedes $u>0$

(3.5) $\frac{q(x-u)}{R(x)} \searrow 0$ *strebt für* $u \leq x \nearrow \infty$.

Beweis. Ersichtlich genügt es zu zeigen, daß (3.1) erfüllt ist. Auf Grund von (3.4) und (3.5) ist aber die in (3.1) auftretende Funktion von x in $x \geq t$ ebenfalls monoton fallend, da die Differenz ihrer Werte für $x=x'$ und $x=x''>x'>t\geq u$

$$= \int_0^t P(u) \left[\frac{q(x'-u)}{R(x')} - \frac{q(x''-u)}{R(x'')}\right] du \geq 0$$

ist. Für festes $t>0$ und $T>t$ ist daher ihre Gesamtschwankung über $t \leq x \leq T$

$$= \frac{1}{R(t)} \int_0^t P(u)\, q(t-u)\, du - \frac{1}{R(T)} \int_0^t P(u)\, q(T-u)\, du.$$

Da hier der Minuend $=1$, der Subtrahend ≥ 0 (aber ≤ 1) ist, ist (3.1) mit $M=1$ erfüllt.

Anwendung. Für

$$p(t) = \frac{t^{k-1}}{\Gamma(k)}, \quad (k>0), \quad \text{und} \quad q(t) = \frac{t^{d-1}}{\Gamma(d)}, \quad (0<d<1),$$

wird, wenn noch $k+d=l$ gesetzt wird,

$$r(t) = p * q = \frac{1}{\Gamma(k)\,\Gamma(d)} \int_0^t (t-u)^{k-1}\, u^{d-1}\, du$$

$$= \frac{t^{l-1}}{\Gamma(k)\,\Gamma(d)} \int_0^1 (1-v)^{k-1}\, v^{d-1}\, dv = \frac{t^{l-1}}{\Gamma(l)}.$$

Die zugehörigen Verfahren N_p und N_r sind ersichtlich die CESÀROschen Verfahren C_k bzw. C_l. Da (3.4) und (3.5) ersichtlich erfüllt sind, letzteres weil abgesehen von konstanten positiven Faktoren

$$\frac{q(x-u)}{R(x)} = \frac{1}{(x-u)^{1-d}\, x^{k+d}}$$

ist und also monoton gegen 0 fällt für wachsendes $x \to \infty$, so ergibt sich aus Satz 4 speziell, daß $|C_k| \leqq |C_l|$ ist für $0 \leqq k \leqq l$.

4. Multiplikationssätze.

Wir benutzen die folgenden Bezeichnungen: Wenn die beiden Integrale

$$(4.1) \qquad \int_0^\infty a(t)\, dt, \qquad (4.2) \qquad \int_0^\infty b(t)\, dt$$

mit $a, b \in S$ vorgelegt sind, setzen wir $a * b = c$ und nennen das Integral

$$(4.3) \qquad \int_0^\infty c(t)\, dt$$

das CAUCHYsche Produkt der beiden ersten. Die Teilintegrale von 0 bis t mögen mit $A(t), B(t), C(t)$ bezeichnet werden und deren NÖRLUND-Mittel, die natürlich jedesmal genauer anzugeben sind, mit $\alpha(x), \beta(x), \gamma(x)$. Die Grenzwerte von diesen sollen schließlich A, B, C genannt werden, wenn sie existieren.

Wenn nun N_p und N_q zwei NÖRLUNDsche Verfahren sind und wenn die Faltung $p * q$ wieder zu $\mathcal{P}$ gehört, so bezeichnen wir sie mit $r(t)$,

$$(4.4) \qquad \int_0^t p(t-u)\, q(u)\, du = r(t),$$

schreiben auch $N_r = N_p * N_q$ (vgl. I, § 5) und setzen endlich die von 0 bis t genommenen Integrale von p, q, r wie bisher beziehentlich $= P(t), Q(t), R(t)$.

Wir können nun den folgenden Satz beweisen, der den klassischen Satz von CESÀRO über die C_1-Summierbarkeit des CAUCHYschen Produktes zweier konvergenter Reihen verallgemeinert:

Satz 5. *Wenn die Verfahren N_p und N_q positiv und regulär sind, wenn* (4.1) *und* (4.2) *beziehentlich N_p- und N_q-summierbar sind zu den Werten A und B, so ist* (4.3) *N_R-summierbar zum Werte $C = A \cdot B$.*

Beweis. Wenn wir $1 * R = \overline{R}$ setzen, so haben wir zu zeigen, daß

$$\bar{\gamma}(x) = \frac{1}{\overline{R}}(R * C) = \frac{1}{\overline{R}(x)} \int_0^x R(x-t)\, C(t)\, dt \to C = A \cdot B \tag{4.5}$$

strebt für $x \to \infty$. Nun ist aber auf Grund der eingangs eingeführten Bezeichnungen

$$P \cdot \alpha = p * A, \quad Q \cdot \beta = q * B$$

und also

$$R * C = 1 * p * q * 1 * a * b = p * A * q * B = (P \cdot \alpha) * (Q \cdot \beta)$$

oder

$$\bar{\gamma}(x) = \int_0^x a(x,t)\, \alpha(t)\, \beta(x-t)\, dt \tag{4.6}$$

mit $a(x,t) = \frac{P(t)\, Q(x-t)}{\overline{R}(x)}$ für $0 < t \leq x$ und $= 0$ für $t > x$.

Wir zeigen nun zunächst, daß die Integraltransformation mit diesem Kern $a(x,t)$ regulär ist. Von den drei Bedingungen, die dafür notwendig und hinreichend sind, sind die Norm- und die Zeilenbedingung offenbar erfüllt, weil $a(x,t) \geq 0$ ist und $\int_0^\infty a(x,t)\, dt = 1$ für alle $x > 0$. Die dritte (Spalten-) Bedingung fordert, daß für jedes feste $c > 0$

$$\frac{1}{\overline{R}(x)} \int_0^c P(t)\, Q(x-t)\, dt \to 0 \quad \text{strebt für} \quad x \to \infty. \tag{4.7}$$

Nun ist aber

$$\int_0^c P(t)\, Q(x-t)\, dt \leq P(c) \int_0^c Q(x-t)\, dt$$

und

$$\overline{R}(x) = \int_0^x P(x-t)\, Q(t)\, dt \geq \int_0^{x-c} \geq P(c) \int_0^{x-c} Q(t)\, dt.$$

Daher ist die linke Seite von (4.7)

$$\leq \int_0^c Q(x-c)\, dt \Big/ \int_0^{x-c} Q(t)\, dt = \{\overline{Q}(x) - \overline{Q}(x-c)\} / \overline{Q}(x-c),$$

wenn $\int_0^t Q(u)\, du = \overline{Q}(t)$ gesetzt wird. Nun ist aber das Verfahren N_Q nach I, Satz 2, Bemerkung 3 regulär und folglich [vgl. I, (2.5)] strebt der letzte Quotient gegen 0 für $x \to \infty$. Das beweist (4.7) und damit die Regularität des Kernes $a(x,t)$. Vertauscht man p und q, so ergibt sich ebenso die Regularität des Kernes $a(x, x-t)$. Unser Satz ist daher eine unmittelbare Folgerung aus dem folgenden

Lemma. *Sind die Kerne $a(x,t)$ und $a(x, x-t)$ beide positiv und regulär, so folgt aus $\alpha(t) \to A$ und $\beta(t) \to B$, daß die durch* (4.6) *erklärte Funktion*

$$\bar{\gamma}(x) \to A \cdot B \quad \textit{strebt für} \quad x \to \infty.$$

Der Beweis ergibt sich, indem wir (4.6) als eine Integraltransformation mit dem Kern

$$\bar{a}(x,t) = a(x,t)\,\beta(x-t)$$

von $\alpha(t)$ in $\bar{\gamma}(x)$ betrachten. Diese Transformation ist nämlich multiplikativ mit dem Faktor B. Denn die Norm- und die Spaltenbedingung sind offenbar erfüllt, da $a(x,t)$ positiv und regulär und $\beta(t)$ beschränkt ist. Die Zeilenbedingung aber liefert den Faktor B, denn es strebt für $x \to \infty$

$$\int_0^\infty \bar{a}(x,t)\,dt = \int_0^\infty a(x, x-t)\,\beta(t)\,dt \to B,$$

weil $a(x, x-t)$ regulär ist und $\beta(t) \to B$ strebt. Wenn also $\alpha(t) \to A$ strebt, so strebt $\bar{\gamma}(x) \to AB$.

Unter einer geeigneten zusätzlichen Voraussetzung läßt sich sogar die N_r-Summierbarkeit des Produktintegrals beweisen:

Satz 6. *Wenn unter denselben Voraussetzungen wie in Satz 5 noch*

$$k(x) = \frac{1}{q(x)}\,(q * b)_x = O(1) \tag{4.8}$$

ist, so ist

$$N_r\text{-}\int_0^\infty c(t)\,dt = AB. \tag{4.9}$$

Wir zeigen zunächst, daß Satz 5 in diesem Satz 6 enthalten ist. Denn wenn wir in Satz 6 N_q und N_r durch N_Q bzw. N_R ersetzen, so wird aus (4.8) die Bedingung

$$\frac{1}{Q(x)}\,(Q * b)_x = \frac{1}{Q(x)}\,(q * B)_x = \beta(x) = O(1) \tag{4.8'}$$

und Satz 6 besagt dann: Wenn N_p und N_Q positiv und regulär sind und wenn überdies $\beta(x) = O(1)$ ist, so folgt aus $N_p\text{-}\int_0^\infty a(t)\,dt = A$ und $N_Q\text{-}\int_0^\infty b(t)\,dt = B$, daß $N_R\text{-}\int_0^\infty c(t)\,dt = AB$ ist. Und hierin ist offenbar Satz 5 enthalten.

Beweis von Satz 6. Wir haben zu zeigen, daß für $x \to \infty$ das N_r-Mittel

$$\gamma(x) = \frac{1}{R(x)}\,(r * C) = \frac{1}{R}\,(p * q * 1 * a * b) \to AB \tag{4.10}$$

strebt. Nun ist aber, da die Faltung kommutativ und assoziativ ist, $p * 1 * a = p * A = P \cdot \alpha$, $q * b = k \cdot q$, also

$$\gamma(x) = \frac{1}{R(x)} \int_0^x P(t)\,\alpha(t)\,q(x-t)\,k(x-t)\,dt = \int_0^x a(x,t)\,k(x-t)\,\alpha(t)\,dt \tag{4.11}$$

$$\text{mit} \quad a(x,t) = \frac{P(t)\,q(x-t)}{R(x)}.$$

Die Integraltransformation mit diesem Kern $a(x,t)$ ist aber (vgl. den Beweis von Satz 5) wieder regulär. Denn die Norm- und die Zeilenbedingung

sind offenbar erfüllt und die Spaltenbedingung bedeutet hier, daß für jedes $c>0$

$$\frac{1}{R(x)}\int_0^c P(t)\,q(x-t)\,dt \to 0$$

strebt für $x\to\infty$. Da aber der Ausdruck links

$$\leqq \frac{P(c)\int_0^c q(x-t)\,dt}{P(c)\int_0^{x-c} q(t)\,dt} = \frac{Q(x)-Q(x-c)}{Q(x-c)}$$

ist und dieser letzte Quotient, da N_q regulär ist, gegen 0 strebt, so ist auch die vorangehende Relation richtig, also $a(x,t)$ regulär. Ganz ähnlich erkennt man, daß die Transformation mit dem Kern $a(x,t)\cdot k(x-t)$ multiplikativ ist mit dem Faktor B. Denn zunächst ist

$$a(x,t)\,k(x-t) = \frac{P(t)\,q(x-t)}{R(x)}\cdot\frac{(q*b)_{x-t}}{q(x-t)} = \frac{P(t)\cdot(q*b)_{x-t}}{R(x)}$$

und daher wegen (4.8) mit einem passenden $K>0$

$$|a(x,t)\,k(x-t)| \leqq K\,\frac{P(t)\,q(x-t)}{R(x)}.$$

Hieraus aber und auf Grund der für den Kern $a(x,t)$ durchgeführten Rechnung ergibt sich, daß der neue Kern die Norm- und die Spaltenbedingung für seine Regularität erfüllt. An Stelle der Zeilenbedingung gilt aber

$$\int_0^x a(x,t)\,k(x-t)\,dt \to B \quad \text{für} \quad x\to\infty,$$

denn die linke Seite dieser Relation ist (s. oben)

$$= \frac{1}{R(x)}\int_0^x P(t)\,(q*b)_{x-t}\,dt = \frac{1}{R(x)}\,(P*q*b)_x$$

$$= \frac{1}{R(x)}\,(p*q*B)_x = \frac{1}{R(x)}\int_0^x r(x-t)\,B(t)\,dt.$$

Dies letzte Mittel strebt aber $\to B$, weil $N_q \subseteq N_r$ gilt (s. I, Satz 2). Also ist (4.11) eine multiplikative Transformation von $\alpha(t)$ in $\gamma(x)$ mit dem Faktor B. Aus $\alpha(t)\to A$ folgt daher $\gamma(x)\to AB$, was den Beweis von Satz 6 vollendet.

War Satz 5 eine Verallgemeinerung des Multiplikationssatzes von CESÀRO, so ist der folgende eine solche des ebenso klassischen Satzes von MERTENS:

Satz 7. *Sind N_p und N_q wieder positiv und regulär und ist* (4.1) *absolut N_p-summierbar zum Werte A und* (4.2) *N_q-summierbar zum Werte B, so ist* (4.3) *N_r-summierbar zum Werte $C=AB$.*

Beweis. Wie in (4.10) haben wir zunächst

$$R \cdot \gamma = r * C = 1 * p * q * a * b,$$

$$1 * (R \cdot \gamma) = p * A * q * B = (P \cdot \alpha) * (Q \cdot \beta)$$

und folglich, wenn wir zur Abkürzung $P \cdot \alpha = f$ und $Q \cdot \beta = g$ setzen, wegen der Stetigkeit von R und γ

$$\gamma(x) = \frac{1}{R(x)} [f * g]'_x. \tag{4.12}$$

Nun sind die vier Funktionen P, α, Q, β und also auch f und g stetig für $x > 0$ und verschwinden für $x \to 0$. Das ist selbstverständlich für P und Q, sowie für A und B. Für $\alpha(x)$ und $x > 0$ folgt die Stetigkeit aus der von $P(x)$ und $\int_0^x p(x-t) A(t)\, dt$ (für die letztere vgl. I, § 3.1). Das letzte Integral ist aber $= \mu \cdot \int_0^x p(t)\, dt = \mu \cdot P(x)$, wenn mit μ ein passender Mittelwert von $A(t)$ in $0 \leq t \leq x$ bezeichnet wird. Daher strebt $\alpha(x) = \mu \to A(0) = 0$ für $x \to 0$. Ebenso erkennt man, daß $\beta(x)$ für $x > 0$ stetig ist und für $x \to 0$ verschwindet.

Nach I, § 3.3 ist daher $(f * g)' = f' * g$, wofern f' zur Klasse $\mathcal{R}$ gehört[8]). Es ist aber, da $P(x)$ ein Integral und $\alpha(x)$ nach Voraussetzung von beschränkter Schwankung ist, fast überall

$$f' = (P \cdot \alpha)' = p\alpha + P\alpha'.$$

Also gilt $f' \in \mathcal{R}$, wofern $\alpha' \in \mathcal{R}$ gilt. Das letztere ist aber gewiß der Fall, weil nach Voraussetzung $\int_0^\infty |d\alpha(t)|$ und folglich auch $\int_0^\infty |\alpha'(t)|\, dt$ endlich ist.

Folglich ist $\gamma(x) = \gamma_1(x) + \gamma_2(x)$ mit

$$\begin{cases} \gamma_1(x) = \dfrac{1}{R(x)} \displaystyle\int_0^x p(x-t)\, \alpha(x-t)\, Q(t)\, \beta(t)\, dt \\ \gamma_2(x) = \dfrac{1}{R(x)} \displaystyle\int_0^x P(x-t)\, \alpha'(x-t)\, Q(t)\, \beta(t)\, dt. \end{cases} \tag{4.13}$$

Diese Darstellungen, als Transformationen von $\beta(t)$ in $\gamma_1(x)$ bzw. $\gamma_2(x)$ betrachtet, sind nun nulltreu. Um dies zu erkennen, genügt es zu zeigen, daß die Kerne die Norm- und die Spaltenbedingung in I, § 2 erfüllen. Für den ersten Kern bedeutet die Normbedingung, da $\alpha(t)$ beschränkt ist, daß

$$\int_0^x p(x-t)\, Q(t)\, dt = O(R(x))$$

[8]) In I, § 3 war $\mathcal{R}$ als die Klasse aller derjenigen für $t > 0$ erklärten reellwertigen Funktionen $f(t)$ definiert, die dieselben Voraussetzungen wie die Gewichtsfunktionen aus $\mathcal{P}$ erfüllen, nur daß $\int_0^x f(t)\, dt$ für $x > 0$ nicht immer $\neq 0$ zu sein braucht.

sein soll. Das ist aber trivialerweise der Fall, da die linke Seite $= R(x)$ ist. Für den zweiten Kern bedeutet sie, daß

$$\int_0^x P(x-t)\,Q(t)\cdot|\alpha'(x-t)|\,dt = O(R(x))$$

sein soll. Doch da für $0 \leq t \leq x$

$$P(x-t)\,Q(t) = P(x-t)\int_0^t q(u)\,du \leq \int_0^t P(x-u)\,q(u)\,du \leq R(x) \tag{4.14}$$

ist, so ist die linke Seite der vorigen Relation in der Tat

$$\leq R(x)\cdot\int_0^x |\alpha'(t)|\,dt \leq R(x)\cdot\int_0^\infty |\alpha'(t)|\,dt = O(R(x)).$$

Die Normbedingung ist also für jeden der Kerne erfüllt. Die Spaltenbedingung verlangt für den ersten Kern, daß für jedes $c>0$

$$\frac{1}{R(x)}\int_0^c Q(t)\,p(x-t)\,dt \to 0$$

strebt. Der Ausdruck zur Linken ist aber, ähnlich wie bei den Schlüssen zu (4.11)

$$\leq \frac{\int_0^c Q(t)\,p(x-t)\,dt}{\int_0^{x-c} p(t)\,Q(x-t)\,dt} \leq \frac{Q(c)}{Q(c)}\cdot\frac{P(x)-P(x-c)}{P(x-c)}$$

und strebt daher gegen 0. Und für den zweiten Kern hat man bei Benutzung von (4.14)

$$\frac{1}{R(x)}\int_0^c P(x-t)\,Q(t)\cdot|\alpha'(x-t)|\,dt \leq \int_0^c |\alpha'(x-t)|\,dt = \int_{x-c}^x |\alpha'(t)|\,dt \to 0.$$

Die beiden Transformationen (4.13) sind daher nulltreu. Wenn also $\beta(t)\to 0$ strebt, so streben $\gamma_1(x)$ und $\gamma_2(x)$ und folglich auch $\gamma(x)\to 0$. Damit ist Satz 7 für den Spezialfall $B=0$ bewiesen.

Ist $B\neq 0$, so betrachten wir statt $b(t)$ die Funktion

$$b_1(t) = \begin{cases} b(t)-B & \text{in} \quad 0\leq t\leq 1 \\ b(t) & \text{in} \quad t>1. \end{cases}$$

Dann ist $\int_0^\infty b_1(t)\,dt$ N_q-summierbar zum Werte 0. Setzen wir also $a * b_1 = c_1$, so ist nach dem eben bewiesenen Spezialfall

$$N_r\text{-}\int_0^\infty c_1(t)\,dt = 0. \tag{4.15}$$

Und hieraus läßt sich die ursprüngliche Behauptung, daß $\int_0^\infty c(t)\,dt$ zum Werte AB N_r-summierbar ist, unschwer ableiten. Es ist nämlich

$$a(x-t)\,b_1(t) = \begin{cases} a(x-t)\,\big(b(t)-B\big), & \text{falls} \quad 0\leq t\leq 1 \\ a(x-t)\,b(t); & \text{falls} \quad t>1 \end{cases}$$

ist, und folglich

$$c_1(x) = \int_0^x a(x-t)\, b_1(t)\, dt =$$

$$= \begin{cases} c(x) - B\int_0^x a(x-t)\, dt = c(x) - B\cdot A(x) & \text{für } 0 \leq x \leq 1, \\ c(x) - \int_0^1 a(x-t)\, dt = c(x) - B\big(A(x) - A(x-1)\big) & \text{für } x > 1 \end{cases}$$

oder also für alle $x \geq 0$

$$c_1(x) = c(x) - B\big(A(x) - A(x-1)\big), \tag{4.16}$$

wenn wir daran festhalten, daß $A(t) = 0$ sein soll für $t < 0$. Setzen wir nun noch $A(t) - A(t-1) = a_1(t)$, so gilt zunächst

$$N_p\text{-}\int_0^\infty a_1(t)\, dt = A. \tag{4.17}$$

Denn es ist

$$\int_0^t a_1(u)\, du = A_1(t) = \int_0^t \big(A(u) - A(u-1)\big)\, du$$

$$= \int_0^t A(u)\, du - \int_0^{t-1} A(u)\, du = \int_{t-1}^t A(u)\, du$$

und folglich ergibt sich (4.17) unmittelbar aus dem folgenden

Hilfssatz. *Wenn das Verfahren N_p positiv und regulär ist und wenn $\int_{t-1}^t s(u)\, du = s_1(t)$ gesetzt wird, so folgt aus N_p-lim $s(t) = s$, daß auch N_p-lim $s_1(t) = s$ ist.*

Beweis. Für die Gewichtsfunktion

$$q_1(t) = \begin{cases} 1 & \text{in } 0 \leq t \leq 1 \\ 0 & \text{in } t > 1 \end{cases}$$

ist

$$q_1 * s = \int_{x-1}^x s(t)\, dt = s_1(x), \tag{4.18}$$

$$r_1 = q_1 * p = \int_{x-1}^x p(t)\, dt = P(x) - P(x-1)$$

und

$$R_1(x) = \int_0^x r_1(t)\, dt = \int_{x-1}^x P(t)\, dt. \tag{4.19}$$

Da nun N_{q_1} offenbar positiv und regulär ist und N_p nach Voraussetzung ebenfalls, so besteht nach I, § 2, Satz 2 die Inklusion $N_p \subseteq N_{r_1}$, d.h. aus N_p-lim $s(t) = s$ folgt N_{r_1}-lim $s(t) = s$. Es ist aber $r_1 * s = p * q_1 * s = p * s_1$, letzteres nach (4.18). Es strebt also

$$\frac{1}{R_1(x)} \int_0^x p(x-t)\, s_1(t)\, dt \to s.$$

Nach (4.19) ist aber $R_1(x) = P(\xi)$ für ein geeignetes ξ zwischen $x-1$ und x und also asymptotisch gleich $P(x)$. Die letzte Beziehung liefert daher die Behauptung des Hilfssatzes und vollendet dadurch den Beweis von (4.17).

Nach (4.16) ist nun

$$c(t) = c_1(t) + B\,a_1(t)$$

und folglich unter Berücksichtigung von (4.15)

$$N_r\text{-}\int_0^\infty c(t)\,dt = B \cdot \left\{ N_r\text{-}\int_0^\infty a_1(t)\,dt \right\}.$$

Wegen (4.17) und $N_p \subseteq N_r$ ist endlich der Wert der in der letzten Beziehung rechts stehenden Klammer $= A$, so daß diese Beziehung die Behauptung des Satzes 7 darstellt.

Literatur.

[1] KNOPP, K., and B. VANDERBURG: Functional NÖRLUND-Methods. I. Rend. Circ. Mat. Palermo (2) 4, Fasc. 1 (1955). — [2] MEARS, FL. M.: Some Multiplication Theorems for NÖRLUND Means. Bull. Amer. Math. Soc. (2) 41, 875—880 (1935). — [3] MEARS, FL. M.: Absolute Regularity and the NÖRLUND Mean. Ann. of Math. (2) 38, 594—601 (1937). — [4] MEARS, FL. M.: The Inverse NÖRLUND Mean. Ann. of Math. (2) 44, 401—410 (1943). — [5] MEARS, FL. M.: NÖRLUND Summability of CAUCHY Products. Ann. of Math. (2) 46, 563—566 (1945). — [6] SUNOUCHI, G., and T. TSUCHIKURA: Absolute Regularity for Convergent Integrals. Tôhoku Math. J. (2) 4, 153—156 (1952).

University of Cincinnati.

Universität Tübingen.

(Eingegangen am 22. Dezember 1954.)

DIEUDONNÉ, J.
Math. Zeitschr. Bd. 63, S. 53—75 (1955)

Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p>0$. III.

A la mémoire du grand algébriste I. SCHUR.

Par

JEAN DIEUDONNÉ.

1. Introduction.

Ce travail fait suite à plusieurs articles antérieurs ([2], [3], [5]), dont nous conservons en principe la terminologie et les notations; il se rattache plus particulièrement à [5], étant exclusivement consacré à l'étude de certains groupes de Lie *abéliens*. Nous y développons notamment une méthode qui permet de démontrer l'*existence* de types très nombreux de groupes de Lie abéliens, deux à deux non isomorphes; parmi ces types figurent les groupes à un paramètre I_r dont l'hyperalgèbre est définie dans [5], et dont nous n'avions pu alors établir l'existence. La méthode dont il s'agit repose sur une étude préalable des *groupes de type hyperexponentiel*, dont la structure est particulièrement simple et qui jouent le rôle d'une sorte de «revêtement universel» pour les groupes que nous étudions; en outre, leur théorie présente un intérêt par elle-même, car elle permet, par exemple, de faire l'étude complète des *groupes additifs de Witt*, du point de vue «infinitésimal».

2. Séries formelles à une infinité de variables graduées.

Il s'avère nécessaire, pour notre objet, de généraliser la notion de groupe de Lie formel, en l'étendant à des séries formelles à une infinité de variables. De façon précise, considérons une suite infinie d'indéterminées $x_0, x_1, \ldots, x_i, \ldots$, et supposons qu'à chacune d'elles soit attaché un entier $\pi_i > 0$, son *poids*, soumis à la condition que *la suite* (π_i) *est strictement croissante et tend vers* $+\infty$; on forme comme d'ordinaire des *monômes*, produits de puissances d'un nombre *fini* des x_i, le poids du produit de deux monômes étant la somme de leurs poids. En raison de l'hypothèse sur la suite (π_i), il n'y a qu'un nombre *fini* de monômes de poids inférieur à un entier donné. Pour un système $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots)$ d'entiers ≥ 0, nuls sauf un nombre fini d'entre eux, nous désignerons encore par x_α le monôme $\prod_{i=0}^{\infty} x_i^{\alpha_i}$; nous désignerons par $\pi(\alpha)$ le poids $\sum_i \alpha_i \pi_i$ de ce monôme; sa hauteur $h(\alpha)$ sera encore le plus petit nombre h tel que $\alpha_i < p^{h+1}$ pour tout indice i; les notations $|\alpha|$ et $\alpha!$ auront le même sens que dans [2, n° 5].

Cela étant, les séries formelles en $x_0, \ldots, x_i, \ldots$ à coefficients dans un corps K sont encore les combinaisons linéaires infinies $\sum_\alpha c_\alpha x_\alpha$, où $c_\alpha \in K$, que l'on ajoute et multiplie comme d'ordinaire [les poids n'interviennent pas ici, mais seulement le fait que pour tout α, il n'y a qu'un nombre fini de couples (β, γ) tels que $\alpha = \beta + \gamma$]. La substitution aux x_i de séries formelles sans terme constant, par rapport à des variables graduées $y_0, y_1, \ldots, y_i, \ldots$, ne peut plus ici se faire sans restriction; elle a toutefois toujours un sens si le *poids minimum* de la série f_i que l'on substitue à x_i (c'est-à-dire le plus petit des poids de ses termes $\neq 0$) *tend vers* $+\infty$ *avec* i: il n'y a alors qu'un nombre fini de monômes x_α qui, lorsqu'on y substitue les f_i aux x_i, fournissent des termes de poids donné.

Nous pouvons alors étendre la définition d'un groupe de Lie formel; il doit y être sous-entendu que, dans tous les systèmes d'indéterminées (x_i), (y_i), etc. que l'on considère, le poids de la variable d'indice i est toujours le *même* entier π_i. La donnée d'un groupe de Lie formel G à une infinité de variables (ou «de dimension infinie») consiste alors en la donnée d'une suite (φ_i) de séries formelles par rapport à deux systèmes (x_i), (y_i) de variables graduées, avec les conditions suivantes: 1° chaque φ_i est de poids minimum π_i; 2° $\varphi_i(x_0, \ldots, x_i, \ldots; 0, \ldots, 0, \ldots) = x_i$, $\varphi_i(0, \ldots, 0, \ldots; y_0, \ldots, y_i, \ldots) = y_i$; 3° $\varphi_i(\varphi(x, y), z) = \varphi_i(x, \varphi(y, z))$ pour tout i, avec les conventions habituelles (les substitutions ayant un sens grâce à la condition 1°). On a alors

$$\varphi_i(x, y) = x_i + y_i + \psi_i(x, y)$$

où ψ_i est une série formelle dont chaque monôme contient au moins un x_j et au moins un y_j, et qui est de poids minimum $\geq \pi_i$. On peut encore démontrer l'existence d'un «inverse», c'est-à-dire de séries $\vartheta_i(x) = -x_i + \omega_i(x)$, où ω_i est de poids minimum $> \pi_i$, telles que $\varphi_i(x, \vartheta(x)) = 0$ pour tout i: il faut ici remplacer la récurrence classique sur le degré des termes des ω_i par une récurrence sur les poids.

Un *homomorphisme* d'un groupe formel G dans un groupe formel G', lorsque G et G' sont de dimension infinie, est un système $u = (u_i)$ de séries formelles $u_i(x_0, \ldots, x_j, \ldots)$ telles que le poids minimum de u_i *tende vers* $+\infty$ *avec* i, et que l'on ait $\varphi'(u(x), u(y)) = u(\varphi(x, y))$. Les notions d'isomorphisme et de «changement de variables» se définissent de même. On a aussi des définitions analogues lorsque l'un des deux groupes G, G' est de dimension finie (lorsque c'est G', il n'y a naturellement plus aucune restriction sur les u_i).

3. Groupes récurrents.

Soit $\mathfrak{o}$ l'anneau des séries formelles par rapport à une infinité de variables graduées $x_0, \ldots, x_i, \ldots$. La notion générale de *semi-dérivation* se définit comme dans [2], ainsi que les semi-dérivations D_α: si on observe que $D_\alpha x_\beta$ est nul ou un monôme de poids $\pi(\beta) - \pi(\alpha)$, on peut en effet définir $D_\alpha\left(\sum_\beta c_\beta x_\beta\right)$ comme $\sum_\beta c_\beta (D_\alpha x_\beta)$ (avec $c_\beta \in K$), cette série formelle ayant un sens. Un

opérateur différentiel D sur $\mathfrak{o}$ pourra cette fois être défini comme une somme infinie $\sum_\alpha v_\alpha D_\alpha$, où les v_α appartiennent à un anneau de séries formelles contenant $\mathfrak{o}$, pourvu que le poids minimum de v_α *augmente indéfiniment avec* $\pi(\alpha)$: la série $\sum_\alpha v_\alpha D_\alpha f$ est en effet définie dans ce cas pour toute $f \in \mathfrak{o}$; en particulier, on a la *formule de Taylor*

$$f(x+z) = \sum_\alpha \frac{1}{\alpha!} z_\alpha D_\alpha f$$

puisque cette formule est vraie pour tout monôme et que z_α est de poids $\pi(\alpha)$.

Nous ne rechercherons pas les conditions les plus générales dans lesquelles une «théorie de Lie» est possible pour les groupes de dimension infinie, et allons nous borner aux seuls groupes qui interviendront par la suite, que nous appellerons *groupes récurrents*. Un tel groupe G est caractérisé par la condition que, pour tout indice i, $\varphi_i(x_0, \ldots, x_j, \ldots; x_0, \ldots, y_j, \ldots)$ *ne contient pas les variables* x_j, y_j *d'indice* $j > i$; cela entraîne que les équations $x_0 = \cdots = x_{i-1} = 0$ définissent un *sous-groupe* (récurrent) G_i' de G; d'autre part, les i séries formelles $\varphi_j(x_0, \ldots, x_{i-1}, 0, \ldots; y_0, \ldots, y_{i-1}, 0, \ldots)$ $(0 \leq j \leq i-1)$ définissent un groupe G_i de dimension i, et le système $u^{(i)} = (u_j^{(i)})$ de séries formelles $u_j^{(i)}(x) = x_j$ pour $j < i$ définit un homomorphisme de G sur G_i, de noyau G_i'. Etant donnés deux groupes récurrents $G, \overline{G}$, un *homomorphisme récurrent* u de G dans $\overline{G}$ est par définition tel que $u_i(x)$ *ne contienne que les variables* x_j *d'indice* $j \leq i$; sa restriction à G_i' est un homomorphisme (récurrent) de ce sous-groupe sur $\overline{G}_i'$ pour tout i, et les i premières séries u_j définissent un homomorphisme de G_i dans $\overline{G}_i$.

La définition des opérateurs de translation L_y et R_y, et des opérateurs différentiels *invariants à gauche* se fait comme dans [2, n^{os} 6 et 7] pour un groupe quelconque de dimension infinie. Si Z_α est l'opérateur différentiel invariant tel que $Z_\alpha(e) = \frac{1}{\alpha!} D_\alpha(e)$, $Z_\alpha f$ est le coefficient de y_α dans le développement de $f(x\,y)$ suivant les puissances des y_i, et on a

$$Z_\alpha = \frac{(1+w_\alpha)}{\alpha!} D_\alpha + \sum_\beta w_\beta D_\beta \tag{1}$$

où w_α et les w_β sont des séries sans terme constant, et la sommation est étendue aux β tels que $|\beta| \leq |\alpha|$ et $\beta \neq \alpha$; cette sommation est ici infinie, mais garde un sens car, en vertu de la formule de Taylor et de l'hypothèse sur les φ_i, le poids minimum de w_β augmente indéfiniment avec $\pi(\beta)$.

Supposons désormais que le groupe G soit *récurrent*; alors on a en particulier

$$X_{hi} = (1 + w_{hi}^{(i)})\, D_{hi} + \sum_{j>i} w_{hj}^{(i)} D_{hj} + \sum_\beta w_\beta^{(i)} D_\beta \tag{2}$$

où toutes les séries $w^{(i)}$ sont sans terme constant, les $w_{hj}^{(i)}$ $(j \geq i)$ appartiennent à $\mathfrak{o}_h$, et les D_β sont des semi-dérivations *spéciales* de hauteur h. En outre, le poids minimum de $w_{hj}^{(i)}$ est $\geq p^h(\pi_j - \pi_i) \geq p^h(j-i)$. Avec ces semi-dérivations invariantes, on forme les «monômes» X_α, qu'il est plus

commode ici de prendre dans l'ordre inverse de celui choisi dans [2, n° 8]: si $\alpha=(\alpha_0, \ldots, \alpha_i, 0, 0, \ldots)$ et $\alpha_j=\sum_{h=0}^{r} \nu_{hj}\, p^h$, on pose

$$X_\alpha = X_{r0}^{\nu_{r0}} \ldots X_{ri}^{\nu_{ri}} X_{r-1,0}^{\nu_{r-1,0}} \ldots X_{r-1,i}^{\nu_{r-1,i}} \ldots X_{00}^{\nu_{00}} \ldots X_{0i}^{\nu_{0i}}.$$

Le point essentiel consiste à préciser les propriétés d'«indépendance linéaire» des D_α et des X_α. En premier lieu, une relation

$$\sum_\mu u_\mu D_\mu = 0 \tag{3}$$

où les u_μ appartiennent à un anneau de séries formelles contenant $\mathfrak{o}$, et sont telles que le poids minimum de u_μ augmente indéfiniment avec $\pi(\mu)$, entraîne $u_\mu=0$ pour tout μ. Pour le voir, on raisonne par l'absurde et on considère dans (3) les termes (en nombre fini) de poids minimum ϱ dans *tous* les u_μ; si ces termes proviennent des séries $u_\alpha, u_\beta, \ldots, u_\lambda$ et si on désigne leur somme par $w_\alpha, w_\beta, \ldots, w_\lambda$ respectivement (dans chacune de ces séries), en écrivant que l'opérateur du premier membre de (3), appliqué successivement aux monômes $x_\alpha, x_\beta, \ldots, x_\lambda$, donne 0, on obtient des séries dont les termes de poids ϱ se réduisent respectivement à $w_\alpha, w_\beta, \ldots, w_\lambda$; ces séries étant nulles, il était contradictoire de supposer $w_\alpha, \ldots, w_\lambda$ non nuls, c'est-à-dire les u_μ non tous nuls.

Remarquons ensuite qu'on peut tirer des équations (2) les formules

$$D_{hi} = (1+u_{hi}^{(i)})\Big(X_{hi} - \sum_\beta w_\beta^{(i)} D_\beta\Big) + \sum_{j>i} u_{hj}^{(i)}\Big(X_{hj} - \sum_\beta w_\beta^{(j)} D_\beta\Big) \tag{4}$$

où les $u_{hj}^{(i)}$ $(j \geq i)$ sont des séries sans terme constant, appartenant à $\mathfrak{o}_h$, et telles que le poids minimum de $u_{hj}^{(i)}$ augmente indéfiniment avec j. En effet, la détermination des $u_{hj}^{(i)}$ se fait par le système récurrent

$$(1+u_{hi}^{(i)})(1+w_{hi}^{(i)}) = 1$$

$$(1+u_{hi}^{(i)})\, w_{hj}^{(i)} + \sum_{k=i+1}^{j-1} u_{hk}^{(i)} w_{hj}^{(k)} + (1+w_{hj}^{(j)})\, u_{hj}^{(i)} = 0 \quad \text{pour} \quad j>i$$

d'où l'en déduit, par récurrence sur $j-i$, que le poids minimum de $u_{hj}^{(i)}$ est $\geq p^h(j-i)$, ce qui établit nos assertions.

L'étape suivante consiste à montrer que pour tout α, on peut écrire

$$D_\alpha = v_\alpha X_\alpha + \sum_\beta v_\beta X_\beta \tag{5}$$

où la sommation est étendue aux β tels que que $|\beta| \leq |\alpha|$, le poids minimum de v_β augmentant indéfiniment avec $\pi(\beta)$; en outre, le terme constant de v_α est égal à 1 (un nombre *fini* d'autres v_β peuvent avoir un terme constant $\neq 0$). Ceci se fait par récurrence sur la hauteur $h(\alpha)$, comme dans [2, n^{os} 8 et 10] (la nouvelle convention adoptée pour les X_α modifiant un peu la démonstration). On a d'ailleurs une formule analogue à (5), avec les mêmes propriétés pour

les coefficients, en y intervertissant les rôles des X_α et des D_α:

$$X_\alpha = t_\alpha D_\alpha + \sum_\beta t_\beta D_\beta. \tag{6}$$

Cette formule est d'ailleurs beaucoup plus élémentaire que (5), et découle aussitôt de (2) et de la définition de X_α, en faisant usage des «formules de Leibniz» pour les D_{hi}. On peut la préciser de la façon suivante. Définissons sur l'ensemble des indices α un ordre total qu'on pourrait appeler «ordre lexicographique inverse par étages»: si $\alpha = (\alpha_0, \ldots, \alpha_i, \ldots)$, $\beta = (\beta_0, \ldots, \beta_i, \ldots)$, α est «avant» β si: 1° ou bien $h(\alpha) > h(\beta)$; 2°, ou bien $h(\alpha) = h(\beta)$, et, si on pose $\alpha_i = \sum_h \mu_{hi} p^h$, $\beta_i = \sum_h \nu_{hi} p^h$, pour le plus grand $h \leq h(\alpha)$ tel que $\mu_{hi} \neq \nu_{hi}$ pour un i au moins, on a $\mu_{hi} > \nu_{hi}$ pour le plus petit indice i pour lequel $\mu_{hi} \neq \nu_{hi}$. On voit alors par récurrence sur $|\alpha|$, en tenant compte de (2), que dans la formule (6), t_α est précisément le «premier» terme non nul, pour l'ordre qui vient d'être défini.

Cette dernière remarque permet enfin de prouver l'indépendance linéaire des X_α, au même sens que pour les D_α. Supposons qu'une relation

$$\sum_\mu u_\mu X_\mu = 0 \tag{7}$$

ait ses coefficients u_μ non nuls; soient encore $w_\alpha, \ldots, w_\lambda$ les termes de poids minimum ϱ dans les u_μ, en nombre fini, et supposons $\alpha, \beta, \ldots, \lambda$ ordonnés suivant l'ordre qui vient d'être décrit. Si on remplace les X_μ dans (7) par leurs expressions (6), on a vu que dans la relation de type (3) entre les D_μ que l'on obtient, tous les coefficients doivent être nuls. Or, les termes de poids minimum dans le coefficient de D_α ont exactement pour somme w_α, puisque α est le «premier» des indices $\alpha, \beta, \ldots, \lambda$ (on se souviendra que le terme constant de t_α est 1). On aurait donc $w_\alpha = 0$, ce qui contredit l'hypothèse. Un raisonnement analogue, mais plus simple, basé sur la formule (1), démontre l'indépendance linéaire des Z_α.

A partir de ces résultats, on peut développer le reste de la théorie exactement comme dans le cas fini. Nous en utiliserons donc librement les résultats, et nous nous bornerons à ajouter quelque remarques tenant à la définition particulière des groupes récurrents. Si dans la formule (1) on remplace les D_μ en utilisant (6) on obtient une relation

$$Z_\alpha = \sum_\mu u_\mu X_\mu \tag{8}$$

et comme Z_α et les X_μ sont des semi-dérivations invariantes, le raisonnement du th. 2 de [2] montre, compte tenu de l'indépendance linéaire des X_μ, que les u_μ sont des *constantes*. Il n'y a par suite qu'un nombre *fini* de ces constantes qui sont $\neq 0$: d'après la formule (1), elles correspondent exactement aux indices μ tels que, dans (5), la série v_μ ait un terme constant $\neq 0$; en particulier, le coefficient de X_α dans (8) est $\frac{1}{\alpha!}$, et les autres coefficients u_μ non nuls correspondent à des indices tels que $|\mu| \leq |\alpha|$. Le fait que G est

un groupe récurrent entraîne en outre que les X_α dont les indices α ont toutes leurs composantes d'indice $k<i$ nulles, forment la base d'une *sous-algèbre* $\mathfrak{G}'_i$ de l'hyperalgèbre $\mathfrak{G}$ de G, savoir l'hyperalgèbre du sous-groupe G'_i [2, n° 18]. Si, dans la formule (8), α a toutes ses composantes d'indice $<i$ nulles, les seuls indices μ tels que $u_\mu \neq 0$ sont donc des indices ayant la même propriété.

Notons enfin que si $\mathfrak{u}^{(i)}$ est l'homomorphisme «canonique» de G sur G_i défini ci-dessus, l'homomorphisme dérivé $(\mathfrak{u}^{(i)})'$ a pour noyau l'*idéal* $\mathfrak{m}_i$ de l'hyperalgèbre de G ayant pour base les X_α (ou les Z_α) tels qu'*une au moins des composantes de α d'indice $\geq i$ soit non nulle*; cela résulte de la formule (13) de [5] et des formules

$$(\mathfrak{u}^{(i)})'(X_{kj}) = X_{kj} \quad \text{pour} \quad j \leq i$$
$$(\mathfrak{u}^{(i)})'(X_{kj}) = 0 \quad \text{pour} \quad j > i.$$

Cette dernière remarque entraîne la propriété suivante: si $\alpha = (\alpha_0, \ldots, \alpha_i, 0, 0, \ldots)$, la série $P_\alpha(\mathfrak{y})$, coefficient de $X_\alpha f$ dans $R_{\mathfrak{y}} f$, *ne contient que les variables x_j d'indice $j \leq i$*; il suffit d'appliquer a $\mathfrak{u}^{(i)}$ les formules (38) et (39) de [2]. Le procédé standard pour obtenir une loi de composition canonique [3] conduit donc à un groupe récurrent (à loi canonique) $\overline{G}$, et définit un isomorphisme récurrent de G sur $\overline{G}$.

4. Les théorèmes d'existence.

Dans [5], nous avons démontré deux propositions générales (prop. 1 et 2) permettant de «remonter» d'une hyperalgèbre de Lie (resp. d'un homomorphisme d'hyperalgèbres) à un groupe (resp. un homomorphisme de groupes). Nous allons leur donner une forme légèrement différente, mieux adaptée aux applications que nous avons en vue. Nous commencerons par nous borner aux groupes de dimension finie.

La base (Z_α) de l'hyperalgèbre de Lie $\mathfrak{G}$ d'un groupe G de dimension n vérifiant la table de multiplication

$$Z_\alpha Z_\beta = \sum_\gamma d_{\alpha\beta\gamma} Z_\gamma, \tag{9}$$

on a les formules

$$Z_\beta x_\alpha = \sum_\gamma d_{\gamma\beta\alpha} x_\gamma \tag{10}$$

et la loi de groupe G s'exprime par les formules

$$\varphi_i(\mathfrak{x}, \mathfrak{y}) = \sum_{\alpha,\beta} d_{\alpha\beta\varepsilon_i} x_\alpha y_\beta \qquad (1 \leq i \leq n). \tag{11}$$

Dans la prop. 1 de [5], on a montré que si, réciproquement, on se donne (abstraitement) une algèbre $\mathfrak{G}$ ayant la table de multiplication (9), elle est l'hyperalgèbre de Lie d'un groupe G *défini* par (11), et on a les relations (10), pourvu que l'on ait les trois conditions suivantes: 1° $d_{0\alpha\beta} = d_{\alpha 0\beta} = \delta_{\alpha\beta}$, $d_{\alpha\beta 0} = 0$ sauf pour $\alpha = \beta = 0$; 2° pour tout $r > 0$, les Z_α de hauteur $h(\alpha) \leq r$ forment la

base d'une sous-algèbre de $\mathfrak{G}$; 3° les $d_{\alpha\beta\gamma}$ vérifient les relations

$$d_{\alpha,\beta,\lambda+\mu} = \sum_{0\leq\gamma\leq\alpha,\,0\leq\delta\leq\beta} d_{\gamma\delta\lambda}\, d_{\alpha-\gamma,\beta-\delta,\mu}\,. \tag{12}$$

pour tous les indices $\alpha, \beta, \lambda, \mu$. Or, cette dernière condition peut se transformer de la façon suivante: considérons l'algèbre $\mathfrak{G}\otimes\mathfrak{G}$, *produit tensoriel* de $\mathfrak{G}$ par elle-même, qui a pour base les produits $Z_\alpha\otimes Z_\beta$. Pour tout indice α, posons

$$Z^0_\alpha = \sum_{0\leq\gamma\leq\alpha} Z_\gamma\otimes Z_{\alpha-\gamma}. \tag{13}$$

On a, d'après (9),

$$\begin{aligned} Z^0_\alpha Z^0_\beta &= \sum_{\gamma,\delta} (Z_\gamma\otimes Z_{\alpha-\gamma})\,(Z_\delta\otimes Z_{\beta-\delta}) \\ &= \sum_{\gamma,\delta} \Big(\sum_\lambda d_{\gamma\delta\lambda}\, Z_\lambda\Big)\otimes\Big(\sum_\mu d_{\alpha-\gamma,\beta-\delta,\mu}\, Z_\mu\Big) \\ &= \sum_{\gamma,\delta,\lambda,\mu} d_{\gamma\delta\lambda}\, d_{\alpha-\gamma,\beta-\delta,\mu}\, Z_\lambda\otimes Z_\mu \end{aligned}$$

et par suite les relations (12) sont *équivalentes* aux relations

$$Z^0_\alpha Z^0_\beta = \sum_\gamma d_{\alpha\beta\gamma}\, Z^0_\gamma. \tag{14}$$

Ces dernières entraînent que les Z^0_α engendrent une *sous-algèbre* $\mathfrak{G}^0$ de $\mathfrak{G}\otimes\mathfrak{G}$ (d'ailleurs isomorphe à $\mathfrak{G}$)[1]. Mais inversement, si cette propriété est vérifiée, toutes les sommes du second membre de (12) qui correspondent à la même valeur ϱ de $\lambda+\mu$ sont égales; en particulier, faisant $\mu=0$, on voit que cette valeur n'est autre que $d_{\alpha\beta\varrho}$ (puisque $d_{\alpha\beta 0}=0$ sauf pour $\alpha=\beta=0$), et on a donc les relations (12). En résumé:

Proposition 1. *Pour qu'une algèbre associative $\mathfrak{G}$, ayant une base (Z_α) dont les indices parcourent N^n, et telle que Z_0 soit l'élément unité de $\mathfrak{G}$, soit l'hyperalgèbre d'un groupe de Lie de dimension n, il faut et il suffit que*: 1° *les Z_α d'indice $\alpha\neq 0$ forment la base d'une sous-algèbre de $\mathfrak{G}$*; 2° *pour tout $r\geq 0$, les Z_α de hauteur $h(\alpha)\leq r$ forment la base d'une sous-algèbre de $\mathfrak{G}$*; 3° *les Z^0_α définis par* (13) *forment la base d'une sous-algèbre $\mathfrak{G}^0$ de $\mathfrak{G}\otimes\mathfrak{G}$.*

Un calcul tout à fait analogue permet de transformer la prop. 2 de [5] en la suivante:

Proposition 2. *Pour qu'un homomorphisme u' de l'hyperalgèbre $\mathfrak{G}$ de G dans l'hyperalgèbre $\overline{\mathfrak{G}}$ de $\overline{G}$, transformant l'élément unité en élément unité, soit dérivé d'un homomorphisme de G dans $\overline{G}$, il faut et il suffit que $u'(\mathfrak{g}_r)\subset\overline{\mathfrak{g}}_r$ pour tout $r\geq 0$ et que $u'\otimes u'$ transforme $\mathfrak{G}^0$ en une sous-algèbre de $\overline{\mathfrak{G}}^0$.*

Les deux propositions précédentes s'étendent sans peine aux groupes récurrents de dimension infinie. Des formules (1), et en se souvenant que $Z_\alpha(e) = \frac{1}{\alpha!}\, D_\alpha(e)$ par définition, on tire aisément les relations (9), où les

[1]) Il est facile de voir que l'isomorphisme de $\mathfrak{G}$ sur $\mathfrak{G}^0$ qui fait correspondre Z^0_α à Z_α n'est autre que le *dérivé* de l'isomorphisme «diagonal» $x\to(x,x)$ de G dans le groupe produit $G\times G$.

$d_{\alpha\beta\gamma} \neq 0$ sont en nombre *fini* et correspondent à des indices γ tels que $|\gamma| \leq |\alpha| + |\beta|$: de façon plus précise, il n'y a qu'un seul de ces indices pour lequel les deux membres de l'inégalité précédente sont égaux, savoir $\gamma = \alpha + \beta$, avec $d_{\alpha,\beta,\alpha+\beta} = \frac{(\alpha+\beta)!}{\alpha!\,\beta!}$; les formules (10) et (11) se démontrent alors comme dans [2, n° 11]. L'hypothèse que G est récurrent s'exprime, en vertu de (11), par les conditions suivantes: on a $d_{\alpha\beta\varepsilon_i} = 0$ si $\pi(\alpha+\beta) < \pi_i$, ou si une des composantes de α ou de β, d'indice $> i$, n'est pas nulle. Inversement, ces conditions, jointes à celles de la proposition 1, assurent que l'algèbre $\mathfrak{G}$ est l'hyperalgèbre d'un groupe récurrent.

Avant de passer aux applications des propositions précédentes, observons en passant que les raisonnements qui y conduisent s'appliquent aussi bien aux groupes de Lie formels sur un corps de caractéristique 0, et conduisent alors aussitôt au «troisième théorème de Lie» pour ces groupes: une fois que l'on sait (th. de Birkhoff-Witt) que l'algèbre enveloppante $\mathfrak{G}$ d'une algèbre de Lie $\mathfrak{g}$ a pour base les X_α, l'hypothèse que $\mathfrak{g}$ est une algèbre de Lie entraîne que les crochets $[X_i^0, X_j^0]$ s'expriment comme combinaisons linéaires des X_k^0, et la prop. 1 montre alors que $\mathfrak{g}$ est l'algèbre de Lie d'un groupe formel.

5. Le groupe hyperexponentiel.

Dans tout ce qui suit, on prend $\pi_i = p^i$. La *série hyperexponentielle* sur un corps K de caractéristique $p > 0$ est une série formelle $\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots)$ par rapport aux variables graduées x_i, qui se définit comme suit [1, n° 4]: on considère, sur le corps $\boldsymbol{Q}$ des rationnels, la série

$$\exp(u_0 + u_1 + \cdots + u_i + \cdots) \tag{15}$$

où les u_i sont des indéterminées, u_i étant affectée du poids p^i; les termes de poids p^i dans cette série ont pour somme un polynôme de la forme $u_i + F_i(u_0, \ldots, u_{i-1})$. Il est immédiat, par récurrence, qu'il existe des polynômes $G_i(x_0, \ldots, x_{i-1})$ à coefficients rationnels, tels que G_i soit isobare de poids p^i, et que le système d'équations $x_i = u_i + F_i(u_0, \ldots, u_{i-1})$ $(i = 0, 1, 2, \ldots)$ se résolve par $u_i = x_i + G_i(x_0, \ldots, x_{i-1})$. Substituant aux u_i ces polynômes dans la série (15), on obtient une série à coefficients rationnels par rapport aux x_i, et on démontre [1, lemme 2] que les coefficients de cette série sont des *entiers p-adiques*. Remplaçant ces coefficients par leurs restes mod. p dans le corps $\boldsymbol{F}_p$, on obtient une série formelle sur ce corps:

$$\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots) = \sum_{k=0}^{\infty} E_k(x_0, \ldots, x_i, \ldots) \tag{16}$$

où on désigne par E_k le polynôme somme des termes de poids k. La propriété fondamentale de cette série est que l'on a

$$\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots)\,\mathrm{Hex}(y_0, y_1, \ldots, y_i, \ldots) = \mathrm{Hex}(z_0, z_1, \ldots, z_i, \ldots) \tag{17}$$

quand on fait dans cette relation la substitution

$$(18) \qquad z_i = \varphi_i(x_0, \ldots, x_i; y_0, \ldots, y_i) = x_i + y_i + \sum_{1<k<p^i} E_k(\boldsymbol{x})\, E_{p^i-k}(\boldsymbol{y})$$

et par suite les formules (18) définissent un groupe récurrent abélien, que nous appellerons *groupe hyperexponentiel* et noterons H ou H_∞. Ses propriétés fondamentales sont résumées dans l'énoncé suivant:

Proposition 3. 1° *Le système des séries formelles* $\boldsymbol{t} = (t_i)$ *telles que*

$$t_0(\boldsymbol{x}) = 0, \qquad t_i(\boldsymbol{x}) = x_{i-1} \qquad (i = 1, 2, \ldots)$$

définit un homomorphisme de H *dans lui-même, appliquant* H *sur* H_0'.

2° *L'hyperalgèbre de* H *est définie par les relations*

$$(19) \qquad X_{hi}^p = X_{h,i+1} \qquad (h = 0, 1, \ldots;\ i = 0, 1, \ldots),$$

et la loi de groupe de H *est canonique.*

1° Il est immédiat, par définition, que l'on a

$$F_i(0, u_0, \ldots, u_{i-2}) = F_{i-1}(u_0, \ldots, u_{i-2})$$

d'où on tire, par récurrence sur i

$$G_i(0, x_0, \ldots, x_{i-2}) = G_{i-1}(x_0, \ldots, x_{i-2})$$

et par suite

$$\mathrm{Hex}(0, x_0, x_1, \ldots) = \mathrm{Hex}(x_0, x_1, \ldots)$$

ce qui, en vertu de (17) et (18), donne

$$(21) \qquad \varphi_i(0, x_0, \ldots, x_{i-1}; 0, y_0, \ldots, y_{i-1}) = \varphi_{i-1}(x_0, \ldots, x_{i-1}; y_0, \ldots, y_{i-1})$$

et démontre notre assertion. La définition de l'homomorphisme $\boldsymbol{t}$ prouve en outre [2, n° 12] que l'on a

$$(22) \qquad \boldsymbol{t}'(X_{hi}) = X_{h,i+1} \qquad (h = 0, 1, \ldots;\ i = 0, 1, \ldots).$$

2° En raison des formules (22), pour démontrer les relations (19), il suffit de considérer le cas $i = 0$. Commençons par calculer X_{00}^p; par définition, on a

$$(23) \qquad X_{00} = D_{00} + \sum_{i=1}^{\infty} w_i D_{0i}$$

où $w_i(\boldsymbol{x})$ est le coefficient de y_0 dans la somme des termes de $\varphi_i(\boldsymbol{x}, \boldsymbol{y})$ qui contiennent y_0 au premier degré et aucun autre y_i. Mais d'après (18), un tel terme ne peut provenir que du produit $E_{p^i-1}(\boldsymbol{x})\, E_1(\boldsymbol{y})$, et comme $E_1(\boldsymbol{y}) = y_0$, on a $w_i(\boldsymbol{x}) = E_{p^i-1}(\boldsymbol{x})$, donc w_i est isobare et de poids $p^i - 1$. Pour $i = 1$, un calcul immédiat montre que $w_1(\boldsymbol{x}) = -x_0^{p-1}$.

Cela étant, dans le calcul des puissances successives X_{00}^t $(2 \leq t \leq p)$ en vue d'obtenir X_{00}^p, seuls importent les coefficients des dérivations D_{0i}, puisque l'on sait d'avance que X_{00}^p est une dérivation. Si $w_i^{(t)}$ est le coefficient de D_{0i}, nous allons voir par récurrence que l'on a $w_1^{(t)} = -(p-1)(p-2)\ldots(p-t+1)\, x_0^{p-t}$, et que pour $i \geq 2$, $w_i^{(t)}$ est nul ou isobare et de poids $p^i - t$.

En effet, supposons qu'il en soit ainsi; alors, pour $j \geq i$, $D_{0j} w_i^{(t)}$ est nul, et pour $j < i$, $D_{0j} w_i^{(t)}$ est nul, ou isobare et de poids $p^i - t - p^j$. Or, $w_i^{(t+1)}$ est égal à $\sum_{j=0}^{\infty} w_j D_{0j} w_i^{(t)}$, donc chacun des termes est nul ou isobare et de poids $p^i - t - 1$, d'où notre assertion (le calcul de $w_1^{(t)}$ étant immédiat). Pour $t = p$, on a donc bien $X_{00}^p = -(p-1)! X_{01} = X_{01}$, puisque l'on peut remplacer x par e dans l'expression de X_{00}^p, et qu'on obtient ainsi D_{01}.

Supposons maintenant les formules (19) prouvées pour $h < r$, et calculons X_{r0}^p. On a évidemment $X_{k\varepsilon_0}^p = X_{k\varepsilon_1}$ en vertu de l'hypothèse, pour $0 \leq k \leq p^r - 1$, et comme les $Z_{k\varepsilon_0}$ sont combinaisons linéaires des $X_{l\varepsilon_0}$ à coefficients dans F_p ($l < p^r$), $Z_{k\varepsilon_0}^p$ s'obtient en remplaçant chaque X_{hi} par $X_{h,i+1}$ ($h < r$) dans l'expression de $Z_{k\varepsilon_0}$, c'est-à-dire en appliquant à $Z_{k\varepsilon_0}$ l'homomorphisme t'; par suite ([5], formule (13)), on a $Z_{k\varepsilon_0}^p = Z_{k\varepsilon_1}$ pour $0 \leq k \leq p^r - 1$. Il en résulte ([5], lemma 2 et formule (6)) que la différence $X_{r0}^p - X_{r1}$ est une *dérivation*. Or, nous allons voir que, pour $r > 0$, il ne figure *aucune dérivation* D_{0j} dans l'expression de X_{ri} à l'aide des D_α (pour tout $i \geq 0$). En effet, le coefficient de D_{0j} serait le coefficient de $y_i^{p^r}$ dans la somme des termes de $\varphi_j(x, y)$ qui contiennent $y_i^{p^r}$ au premier degré et aucun autre y_k. D'après (18), un tel terme ne peut provenir que du produit $E_{p^j - p^{i+r}}(x) E_{p^{i+r}}(y)$ (si $j > i + r$); mais comme $E_{p^{i+r}}(y) = y_{i+r}$, notre assertion est démontrée. A fortiori, il ne peut figurer aucune dérivation dans un produit X_α qui n'est pas un produit de dérivations X_{0j}, et nous voyons donc que $X_{r0}^p - X_{r1} = 0$, ce qui achève de démontrer les formules (19).

Reste à voir que la loi de composition de H est canonique; il nous suffira de prouver que, si X_α n'est pas un des X_{0j}, dans la formule (6), il n'y a aucune série formelle t_{ε_j} dont le terme constant ne soit pas nul. Supposons en effet ce point établi: alors, pour $\alpha \neq \varepsilon_i$, montrons que le second membre de la formule (8) ne peut contenir aucun X_{0j}, ce qui signifiera que la loi de groupe est canonique. S'il en était autrement, en substituant dans (8) à chaque X_μ son expression tirée de (6), puis en remplaçant x par e, on aurait une relation linéaire (à coefficients constants) entre les Z_β, qui devrait donc avoir ses coefficients nuls. Mais comme par hypothèse aucun $Z_{\varepsilon_j} = X_{0j}$ ne peut provenir de l'expression (6) de X_μ si μ n'est pas lui-même de la forme ε_k, les termes en Z_{ε_j} du second membre de (8) ne pourraient provenir que des termes en X_{0j}, et par suite certains auraient dans ce second membre un coefficient non nul, ce qui est absurde si $|\alpha| > 1$.

Pour achever notre démonstration, remarquons que si $h(\alpha) \geq 1$, les séries formelles t_{ε_j} du second membre de (6) sont en fait *nulles*, comme nous l'avons remarqué ci-dessus. On est donc ramené au cas où $X_\alpha = X_{00}^{\alpha_0} X_{01}^{\alpha_1} \dots X_{0k}^{\alpha_k}$, avec $0 \leq \alpha_i \leq p - 1$ pour $0 \leq i \leq k$ et $|\alpha| > 1$. Raisonnons par récurrence sur k, la proposition ayant déjà été établie pour $k = 0$ dans la première partie du raisonnement. Remarquons tout d'abord que pour calculer X_{0i} il faut prendre les coefficients de y_i dans les monômes des $\varphi_j(x, y)$ qui contiennent y_i au premier degré et aucun autre y_k, les multiplier par D_{0j} et en faire la somme. La considération des poids des E_k montre que, dans $\varphi_j(x, y)$, ces monômes

sont ceux du produit $E_{p^i-p^j}(x)E_{p^j}(y)=y_iE_{p^j-p^i}(x)$; autrement dit, on a

$$X_{0i}=D_{0i}+\sum_{j>i}E_{p^j-p^i}(x)\,D_{0j}. \tag{24}$$

Supposons d'abord $\alpha_0>0$; on peut compléter les remarques faites au 1° concernant l'expression de $X_{00}^{\alpha_0}$ en remarquant que dans cette expression, seul le terme $D_{00}^{\alpha_0}$ a un coefficient dont le terme constant est $\neq 0$; il est en effet immédiat par récurrence sur α_0 que dans $X_{00}^{\alpha_0}$, le coefficient de D_{01} est celui (en dehors du terme $D_{00}^{\alpha_0}$) dont le poids minimum est le plus petit, et ce poids est $p-\alpha_0>0$. Si $X_{01}^{\alpha_1}X_{02}^{\alpha_2}\dots X_{0k}^{\alpha_k}=\sum_\beta v_\beta D_\beta$, il ne pourrait donc y avoir dans X_α des coefficients de dérivations D_{0j} dont le terme constant serait $\neq 0$ que si ces coefficients étaient des séries de la forme $D_{00}^{\alpha_0}v_{\alpha_j}$. Or, observons que dans tout monôme d'un polynôme de la forme $E_{ps}(x)$ où figure x_0, l'exposant de x_0 est *multiple de* p, car le poids total de E_{ps} et le poids de chacun des x_i d'indice $i\geq 1$ sont multiples de p. Compte tenu des expressions (24), on en conclut aussitôt que les v_β sont des polynômes en $x_0^p, x_1, \dots, x_k$, et par suite $D_{00}v_\beta=0$ pour tout β. Notre assertion est donc démontrée dans ce cas.

Supposons ensuite que $\alpha_0=0$. Pour trouver les coefficients des D_{0j} dans X_α qui ont un terme constant $\neq 0$, on peut évidemment commencer par remplacer x_0 par 0 dans ces coefficients. Or, si on observe que dans les expressions (24) pour $i\geq 1$, D_{00} ne figure pas, on voit que le nouvel opérateur que l'on obtient après la substitution précédente s'obtient aussi en remplaçant *d'abord* x_0 par 0 dans les coefficients de (24), puis en formant le monôme $X_{01}'^{\alpha_1}\dots X_{0k}'^{\alpha_k}$ à partir des opérateurs X_{0i}' ainsi obtenus. Mais il suit de (18) et (21) que l'on a

$$E_{ps}(0, x_1, \dots, x_i, \dots)=E_s(x_1, \dots, x_i, \dots) \tag{25}$$

comme les considérations de poids le montrent aussitôt. On voit par suite que X_{0i}' se déduit de $X_{0,i-1}$ en remplaçant chacun des indices j qui figurent dans $X_{0,i-1}$ par $j+1$ (aussi bien dans les indéterminées que dans les dérivations). L'hypothèse de récurrence prouve par suite que dans X_α aucun des D_{0j} n'a un coefficient dont le terme constant est $\neq 0$, et la proposition 3 est ainsi complètement démontrée.

6. Groupes de type hyperexponentiel.

Nous dirons qu'un groupe *récurrent* G est *de type hyperexponentiel* s'il satisfait aux trois conditions suivantes:

1° G est *abélien*.

2° L'algèbre de Lie $\mathfrak{g}_0$ de G a pour table de multiplication

$$X_{0i}^p=X_{0,i+1} \qquad (i=0,1,\dots). \tag{26}$$

On notera que les conditions (26) entraînent, en vertu de l'existence de l'homomorphisme de Frobenius p' [2, n° 13], que l'on a nécessairement, pour tout $h\geq 1$

$$X_{hi}^p=X_{h,i+1}+\sum_\lambda a_\lambda X_\lambda \tag{27}$$

où, dans la sommation, tous les indices λ sont de hauteur $< h$ et n'ont aucune composante d'indice $< i$ qui soit non nulle. Cela étant, la troisième condition imposée aux groupes de type hyperexponentiel est la suivante:

3° Dans la formule (27), pour tout indice λ tel que $a_\lambda \neq 0$, il y a *au moins une* composante de λ d'indice $> i$ qui est *non nulle*.

Remarquons que ces trois conditions sont *invariantes* par tout *isomorphisme récurrent* v de G sur $\overline{G}$, tel que $v_i(x) = x_i + \cdots$ pour tout i. Cela est immédiat pour les deux premières, puisque v' laisse invariante l'algèbre de Lie. En outre, en raison de l'hypothèse sur les premiers termes des v_i, on a $(v')^{-1}(\overline{X}_{hi}) = X_{hi} + \sum_\mu b_\mu X_\mu$, où dans la sommation, $h(\mu) < h$ et les composantes d'indice $< i$ des μ sont toutes nulles. On en déduit que $(v')^{-1}(\overline{X}^p_{hi})$ est une combinaison linéaire $\sum_\nu c_\nu X_\nu$, où dans chaque indice ν de coefficient $c_\nu \neq 0$, toutes les composantes d'indice $< i$ sont nulles et une composante au moins d'indice $> i$ n'est pas nulle; cela signifie encore que cet élément appartient à l'idéal $\mathfrak{m}_i$, noyau du dérivé de l'homomorphisme canonique $u^{(i)}$ de G sur G_i (n° 3). Mais comme v est un isomorphisme récurrent, on a $\overline{u}^{(i)} v = v\, u^{(i)}$, en désignant par $\overline{u}^{(i)}$ l'homomorphisme canonique de $\overline{G}$ sur $\overline{G}_i$. Si $\overline{\mathfrak{m}}_i$ est le noyau du dérivé de $\overline{u}^{(i)}$, on a par suite $v'(\mathfrak{m}_i) = \overline{\mathfrak{m}}_i$, d'où $\overline{X}^p_{hi} = v'\left(\sum_\nu c_\nu X_\nu\right) \in \overline{\mathfrak{m}}_i$, ce que nous voulions démontrer.

La définition que nous venons de donner est justifiée par le théorème suivant:

Théorème 1. *Si G est un groupe de type hyperexponentiel, il existe un isomorphisme récurrent u de G sur le groupe hyperexponentiel H, tel que $u_i(x) = x_i + \cdots$ (termes non écrits de degré > 1) pour tout $i \geq 0$.*

D'après les remarques qui précèdent l'énoncé, on peut supposer que la loi de groupe de G est *canonique* (n° 3, *in fine*). Nous allons démontrer le théorème par «récurrence sur la hauteur» suivant un procédé déjà utilisé dans [5]. Nous supposerons donc que l'hyperalgèbre de G vérifie, non seulement (26), mais aussi les relations (19) *pour tout i et pour $h < r$*. Notons d'abord que, d'après la démonstration du th. d'unicité ([3, th. 2]; voir aussi [4, n° 4]), les algèbres $\mathfrak{g}_r$ ayant même table de multiplication pour G et H, et ces deux groupes ayant des lois canoniques (prop. 3), les coefficients $d_{\alpha\beta\varepsilon_i}$ qui définissent les lois de groupe [formule (11)] sont les mêmes pour G et H, lorsque α et β sont *tous deux de hauteur $< r$*; on en conclut que, pour $h(\alpha) < r$, l'expression de Z_α à l'aide des X_λ a les *mêmes coefficients* dans G et dans H. On a donc, pour le groupe G

$$Z^p_{(\alpha_0, \alpha_1, \ldots)} = Z_{(0, \alpha_0, \alpha_1, \ldots)}$$

pour $h(\alpha) < r$, puisque ces relations ont lieu dans H. D'après [5, lemma 2], on en déduit que $X^p_{ri} - X_{r,i+1}$ est une *dérivation*, donc une combinaison linéaire des X_{0j}; d'ailleurs, en vertu de la condition 3° dans la définition d'un groupe de type hyperexponentiel, on a

$$X^p_{ri} - X_{r,i+1} = \sum_{j \geq i+1} \lambda_{ij} X_{0j} \qquad (i = 0, 1, \ldots). \tag{28}$$

Considérons un isomorphisme récurrent v_1 de G sur un groupe $G^{(1)}$, défini par

$$v_j(x) = x_j + \sum_{i \leq j} \mu_{ji}\, x_i^{p^{j-i}} \qquad (j = 0, 1, \ldots).$$

On déduit de ces formules

$$v_1'(X_{hi}) = X_{hi}^{(1)} \quad \text{pour} \quad h < r$$

$$v_1'(X_{ri}) = X_{ri}^{(1)} + \sum_{j \geq i} \mu_{ji} X_{0j}^{(1)}$$

et par suite, en vertu de (28) et de l'hypothèse sur G,

$$(X_{ri}^{(1)})^p - X_{r,i+1}^{(1)} = \sum_{j \geq i+1} (\lambda_{ij} + \mu_{j,i+1} - \mu_{j-1,i}^p)\, X_{0j}^{(1)}. \tag{29}$$

On voit aussitôt qu'on peut déterminer les μ_{ji} de façon que les seconds membres de (29) s'annulent pour tout i: on peut prendre $\mu_{j0} = 0$ pour tout j, et les μ_{ji} pour $i \geq 1$ sont déterminés par un système récurrent ($j = i, i+1, \ldots$). On obtient ainsi un groupe $G^{(1)}$ de type hyperexponentiel satisfaisant aux relations (19) pour tout i, et pour $h \leq r+1$; sa loi de composition n'est peut-être pas canonique, mais le procédé standard [3] fournit un isomorphisme récurrent v_2 de $G^{(1)}$ sur $\overline{G}$ satisfaisant aux mêmes propriétés et ayant une loi canonique. Le processus de récurrence peut donc se poursuivre et conduit finalement (par le «passage à la limite» habituel, cf. [3] et [5]) à un isomorphisme de G sur un groupe H' de type hyperexponentiel, ayant une loi canonique et dont l'hyperalgèbre est définie par (19). En vertu de la prop. 3 et du th. d'unicité [3, th. 2], H' est identique à H et le théorème est démontré.

Corollaire. *Si G est un groupe de type hyperexponentiel, le groupe G_i' est isomorphe à H et le groupe G_i est isomorphe à H_i.*

Cela résulte des propriétés générales des isomorphismes récurrents (n° 3) et du fait que H_i' est isomorphe à H.

7. Les groupes de Witt.

Le *groupe additif de Witt*, que nous noterons W ou W_∞, est, rappelons-le [6], défini de la façon suivante: on considère, sur le corps des nombres rationnels, le système récurrent

$$\begin{cases} z_0 = x_0 + y_0 \\ z_0^p + p\, z_1 = (x_0^p + p\, x_1) + (y_0^p + p\, y_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_0^{p^i} + p\, z_1^{p^{i-1}} \cdots + p^i z_i = (x_0^{p^i} + p\, x_1^{p^{i-1}} + \cdots + p^i x_i) + (y_0^{p^i} + p\, y_1^{p^{i-1}} + \cdots + p^i y_i) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases} \tag{30}$$

dont on démontre que la solution $z_i = L_i(x_0, y_0, x_1, y_1, \ldots, x_i, y_i)$ $(i = 0, 1, \ldots)$ est telle que L_i soit un polynôme à coefficients entiers, isobare et de poids p^i; en réduisant les coefficients mod. p on obtient une loi de groupe récurrent abélien qui définit W.

De même, le «groupe des Einseinheiten» de Witt (noté V ou V_∞) est défini en considérant le système récurrent (sur le corps des rationnels)

$$(31)\quad \begin{cases} 1+p\,z_0 = (1+p\,x_0)(1+p\,y_0) \\ 1+p\,z_0^p+p^2 z_1 = (1+p\,x_0^p+p^2 x_1)(1+p\,y_0^p+p^2 y_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1+p\,z_0^{p^i}+\cdots+p^{i+1}z_i = (1+p\,x_0^{p^i}+\cdots+p^{i+1}x_i)(1+p\,y_0^{p^i}+\cdots+p^{i+1}y_i) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

Sa solution $z_i = M_i(x_0, y_0, x_1, y_1, \ldots, x_i, y_i)$ $(i = 0, 1, \ldots)$ est encore telle que M_i soit un polynôme à coefficients entiers, qui ici n'est plus isobare, mais dont le poids minimum est p^i; en réduisant les coefficients mod. p, on obtient la loi de groupe de V.

Les groupes V et W sont des *groupes de type hyperexponentiel*, donc *isomorphes* à H. La vérification des relations (26) peut en effet se faire exactement comme pour le groupe H (n° 5): il suffit d'observer que l'on a

$$(32)\qquad X_{0i} = D_{0i} + \sum_{j \geq i+1} w_{ij} D_{0j}$$

et qu'en raison de la définition des groupes récurrents, le poids minimum de w_{ij} est $\geq p^j - p^i$; en outre, dans $w_{i,i+1}$ figure le monôme $-x_i^{p-1}$, et les autres monômes contiennent tous au moins un x_k d'indice $k<i$. Reste à prouver la troisième condition de la définition d'un groupe de type hyperexponentiel. Mais comme, pour V et W, la série formelle $\varphi_i(\boldsymbol{x}, \boldsymbol{y})$ est de la forme $x_i + y_i + \psi_i(x_0, y_0, \ldots, x_{i-1}, y_{i-1})$, X_{hi} est somme de D_{hi} et d'une combinaison linéaire de D_β, où dans chaque β, une au moins des composantes d'indice $> i$ est $\neq 0$. Comme $D_{hi}^p = 0$, X_{hi}^p s'exprime comme combinaison linéaire de D_β où les indices β ont la même propriété, et en remplaçant $\boldsymbol{x}$ par $\boldsymbol{e}$, on en conclut que X_{hi}^p est combinaison linéaire des Z_β correspondant à ces indices, ce qui achève le raisonnement.

Le lecteur observera que nous sommes loin d'avoir utilisé la forme particulière des lois des groupes V et W, et nous lui laissons le soin de formuler des conditions générales permettant d'affirmer qu'un groupe récurrent est du type hyperexponentiel. En fait, on peut obtenir des résultats beaucoup plus précis pour le groupe W, et prouver entre autres que son hyperalgèbre a la même table de multiplication (19) que celle de H; les deux groupes ne sont toutefois pas identiques, mais ont entre eux des relations assez remarquables. Comme ces questions sont un peu étrangères à l'objet principal de ce travail, et comportent des démonstrations assez longues, nous les exposerons dans une publication ultérieure.

8. Le groupe universel.

Nous allons définir encore un autre groupe de type hyperexponentiel, mieux adapté que les précédents aux applications que nous avons en vue:

Proposition 4. *Il existe un groupe U de type hyperexponentiel sur le corps $\mathbf{F}_p$ dont l'hyperalgèbre de Lie est définie par les relations* (19), *et dont*

l'opérateur de translation est donné par la formule

$$R_y = \prod_{i=0}^{\infty} \mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^h} X_{hi}, \ldots). \tag{33}$$

Nous allons utiliser la prop. 1. Les relations (19) montrent que, du point de vue abstrait, l'hyperalgèbre du groupe dont nous voulons démontrer l'existence est isomorphe à *l'algèbre des polynômes* P (à coefficients dans $\boldsymbol{F}_p$) par rapport aux indéterminées

$$T_0 = X_{00}, \quad T_1 = X_{10}, \ldots, \quad T_h = X_{h0}, \ldots.$$

On a donc

$$\mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^h} X_{hi}, \ldots) = \sum_{k=0}^{\infty} y_j^k E_k(T_0^{p^j}, T_1^{p^j}, \ldots, T_h^{p^j}, \ldots).$$

Considérons les polynômes

$$E_\alpha(T_0, \ldots, T_h, \ldots) = \prod_{i=0}^{\infty} E_{\alpha_i}(T_0^{p^i}, \ldots, T_h^{p^i}, \ldots) \tag{34}$$

pour tout indice $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots)$ ($\alpha_i = 0$ à partir d'un certain rang). Montrons d'abord que les E_α forment une *base* de l'algèbre de polynômes P. Soit $k = \nu_0 + \nu_1 p + \cdots + \nu_h p^h$ un entier ≥ 0 développé p-adiquement; si on décompose la série (15) en produit des exponentielles $\exp(u_i)$, on voit aussitôt qu'il y a dans la série (15) un terme de poids k égal à $\frac{u_0^{\nu_0} u_1^{\nu_1} \ldots u_h^{\nu_h}}{\nu_0! \, \nu_1! \ldots \nu_h!}$, et que si on ordonne les monômes en les u_i suivant l'ordre lexicographique *inverse* (le «premier» de deux monômes $u_0^{\alpha_0} \ldots u_h^{\alpha_h} \ldots$, $u_0^{\beta_0} \ldots u_h^{\beta_h} \ldots$ étant $u_0^{\alpha_0} \ldots u_h^{\alpha_h} \ldots$ si et seulement si, pour le *dernier* des indices j tels que $\alpha_j \neq \beta_j$, on a $\alpha_j > \beta_j$), le terme $\frac{u_0^{\nu_0} \ldots u_h^{\nu_h}}{\nu_0! \ldots \nu_h!}$ est le «premier» des termes de poids k. On en conclut que, lorsqu'on substitue à chaque u_i le polynôme isobare $x_i + G_i(x_0, \ldots, x_{i-1})$, on obtient, comme «premier» terme de poids k, le monôme $\frac{x_0^{\nu_0} \ldots x_h^{\nu_h}}{\nu_0! \ldots \nu_h!}$; son coefficient n'étant pas multiple de p, c'est aussi ce monôme qui est le «premier» dans $E_k(x_0, \ldots, x_i, \ldots)$. Comme l'ordre lexicographique inverse est compatible avec l'addition dans $\boldsymbol{N}^{(\boldsymbol{N})}$, le «premier» terme de E_α est le produit des «premiers» termes des E_{α_i} dans la formule (34). On en conclut aisément que les E_α sont linéairement indépendants sur $\boldsymbol{F}_p$: il suffit, dans une relation linéaire entre ces polynômes (en nombre fini), de considérer l'ensemble des termes de plus petit poids, dont la somme doit déjà être nulle; on considère ensuite la somme des «premiers» termes des E_α (supposés de même poids), et tout revient à voir que si $\alpha \neq \beta$, les «premiers» termes de E_α et de E_β sont distincts. Mais si on pose $\alpha_i = \lambda_{0i} + \lambda_{1i} p + \cdots + \lambda_{hi} p^h$ pour tous les $\alpha_i \neq 0$, avec $0 \leq \lambda_{ji} \leq p-1$, on peut supposer que h est le même pour tous les α_i (avec éventuellement $\lambda_{hi} = 0$). Si on écrit $m_k = \sum_{i=0}^{\infty} \lambda_{ki} p^i$ ($0 \leq k \leq h$), le «premier» terme de E_α est le monôme $\frac{1}{\alpha!} T_\mu = \frac{1}{\mu!} T_\mu$, où

$$T_\mu = T_0^{m_0} T_1^{m_1} \ldots T_h^{m_h}. \tag{35}$$

Il est clair que les «premiers» termes de E_α et E_β ne peuvent être égaux que si les λ_{hi} sont les mêmes pour $0 \leq k \leq h$, $0 \leq i < +\infty$, ce qui signifie que $\alpha = \beta$.

Reste à montrer que tout monôme (35) est combinaison linéaire des E_α; or, si à partir des m_k on forme les α_i par le procédé inverse de celui décrit ci-dessus, et si on forme $T_\mu - \alpha! E_\alpha$, on obtient un polynôme isobare de *même poids* que T_μ, et dont le «premier» terme est «après» celui de E_α. Comme l'ensemble des monômes *de poids donné* est fini, on est ramené, par récurrence, à démontrer notre assertion pour les «derniers» monômes de poids donné (pour l'ordre lexicographique inverse), c'est-à-dire les monômes $T_0^{m_0}$. Or, pour $0 \leq k < p$, on a $E_k(T_0) = T_0^k/k!$; donc $T_0^{m_0}$ est égal à $\prod_{i=0}^{\infty} E_{\lambda_{0i}}(T_0^{p^i})$ à un facteur près; ceci achève de démontrer que les E_α forment une base de P.

Ceci étant établi, on peut écrire

$$E_\alpha E_\beta = \sum_\gamma d_{\alpha\beta\gamma} E_\gamma \tag{36}$$

et comme E_k ne contient que $T_0, \ldots, T_{k-1}$ si $k < p^h$, les E_α correspondant aux α de hauteur $< h$ engendrent une sous-algèbre de P, savoir l'algèbre P_h des polynômes en $T_0, \ldots, T_{h-1}$. Pour appliquer la prop. 1, il faut considérer le produit tensoriel $P \otimes P$; ce n'est autre comme on sait (à une isomorphie près), que l'algèbre des polynômes par rapport aux deux séries d'indéterminées $T_0, T_1, \ldots$; $U_0, U_1, \ldots$, et elle a pour base (lorsqu'on fait l'identification précédente) les produits $E_\alpha(T_0, \ldots, T_h, \ldots) E_\beta(U_0, \ldots, U_h, \ldots)$. Or, la formule (17) montre que

$$\sum_{j=0}^{k} E_j(T_0, \ldots, T_h, \ldots) E_{k-j}(U_0, \ldots, U_h, \ldots) = E_k(\Phi_0, \Phi_1, \ldots, \Phi_h, \ldots) \tag{37}$$

où on a posé $\Phi_h = \varphi_h(T_0, \ldots, T_h; U_0, \ldots, U_h)$. De cette formule et de (34), on déduit aussitôt (compte tenu de ce que tous les coefficients sont dans $\boldsymbol{F}_p$)

$$E_\alpha(\Phi_0, \Phi_1, \ldots, \Phi_h, \ldots) = \sum_{0 \leq \beta \leq \alpha} E_\beta(T_0, \ldots, T_h, \ldots) E_{\alpha-\beta}(U_0, \ldots, U_h, \ldots). \tag{38}$$

Mais comme (36) est une identité polynomiale, on peut y substituer Φ_h à T_h pour tout h, et on conclut de (38) que les éléments $\sum_{0 \leq \beta \leq \alpha} E_\beta(T_0, \ldots, T_h, \ldots)$ $E_{\alpha-\beta}(U_0, \ldots, U_h, \ldots)$ forment la base d'une sous-algèbre de $P \otimes P$ (identifiée à l'algèbre des polynômes par rapport aux T_h et aux U_h). Enfin, comme E_α est isobare et de poids $\pi(\alpha)$, il est clair que $d_{\alpha\beta\varepsilon_i} = 0$ si $\pi(\alpha+\beta) < p^i$ ou si une des composantes de α (ou de β), d'indice $j > i$, n'est pas nulle, puisqu'alors $\pi(\alpha) \geq p^j$. Toutes les conditions d'application de la prop. 1 (ou plutôt de son extension aux groupes récurrents, cf. n° 4) sont par suite réalisées, et la prop. 4 est démontrée.

Nous dirons que le groupe récurrent U est le *groupe universel* (terminologie qui sera justifiée au n° 9). Une conjecture naturelle serait que les groupes H et U sont identiques; il n'en est malheureusement rien, comme nous allons

le voir sur un exemple. En raison de la prop. 3 et du fait que les tables de multiplication des hyperalgèbres de H et de U sont les *mêmes*, il suffit de montrer que la loi de groupe de U *n'est pas canonique* en général; autrement dit, il nous suffira de donner un exemple d'un opérateur E_α de l'hyperalgèbre de U, tel que $\alpha \neq \varepsilon_i$, et qui contient un terme en X_{0j} dans son expression à l'aide des X_λ. Il suffit de prendre $p=3$, $\alpha=(6, 1, 0, 0, \ldots)$; on trouve facilement $E_6(T_0, T_1, \ldots) = T_0^6 - T_1^2$, d'où $E_\alpha(T_0, T_1, \ldots) = T_0^9 - T_1^2 T_0^3 = X_{02} - X_{10}^2 X_{01}$.

9. Les «groupes quotients» du groupe universel.

Nous avons démontré dans [5, th. 3] qu'un groupe de Lie formel abélien G de dimension finie sur un corps algébriquement clos est produit direct d'un certain nombre de groupes isomorphes au groupe multiplicatif W_1^* (dont le produit constitue ce que nous avons appelé le *noyau* («core») de G), et d'un *groupe radical*, dont l'hyperalgèbre est une *nilalgèbre* (i.e., tous ses éléments sont nilpotents). Nous nous proposons de montrer comment on peut, à partir du groupe universel U, définir des familles de groupes radicaux, qui en sont des *images homomorphes*, et qui contiennent en particulier tous les groupes radicaux à un paramètre I_r ($r \geq 1$) qui (à l'existence près) avaient été caractérisés dans [5] par leur hyperalgèbre.

Avant de développer la méthode donnant ces groupes, remarquons en passant que le groupe multiplicatif lui-même apparaît comme image homomorphe de U, ou mieux ici de H. En effet, par sa définition même, la série hyperexponentielle $\mathrm{Hex}(x_0, x_1, \ldots)$ constitue un homomorphisme du groupe H sur le groupe multiplicatif W_1^*. Une conséquence immédiate de ce fait est la propriété «différentielle» de cette série, qui la rapproche de l'exponentielle classique: pour *tout opérateur différentiel invariant* Δ du groupe H, on a

$$\Delta\big(\mathrm{Hex}(x_0, x_1, \ldots)\big) = \mathrm{Hex}(x_0, x_1, \ldots). \tag{39}$$

En effet, on a par définition $\Delta R_y = R_y \Delta$, et ici

$$R_y\big(\mathrm{Hex}(x_0, x_1, \ldots)\big) = \mathrm{Hex}(y_0, y_1, \ldots)\,\mathrm{Hex}(x_0, x_1, \ldots).$$

Il suffit d'appliquer Δ aux deux membres, puis, tenant compte de $\Delta R_y = R_y \Delta$, de faire $y = e$ dans l'équation obtenue, pour avoir (39).

Comme premiers exemples de groupes radicaux déduits du groupe universel, nous allons considérer des groupes dont l'*algèbre de Lie* $\mathfrak{g}_0$ est *indécomposable*, autrement dit [5, n° 11] admet une base telle que $X_{0i}^p = X_{0,i+1}$ pour $0 \leq i < n-1$, $X_{0,n-1}^p = 0$. Soient m et r deux entiers tels que $0 \leq m \leq n-1$, $r \geq 1$. Nous allons montrer qu'*il existe un groupe* $G_{n,m,r}$ *de dimension* n *dont l'hyperalgèbre a pour table de multiplication*

$$\begin{cases} X_{hi}^p = X_{h,i+1} & (h = 0, 1, \ldots;\ 0 \leq i < n-1) \\ X_{h,n-1}^p = 0 \quad \text{pour} \quad h < r \\ X_{r-h,n-1}^p = X_{hm} \quad \text{pour} \quad h = 0, 1, \ldots. \end{cases} \tag{40}$$

Il semble assez commode de représenter cette table de multiplication sous forme d'un *diagramme*, où les X_{hi} sont rangés en un tableau à double entrée suivant la convention habituelle, et où une flèche joignant deux points signifie que l'extrémité de la flèche est la puissance p-ème de l'origine; le diagramme ci-dessous correspond donc au cas $n=4$, $m=1$, $r=3$:

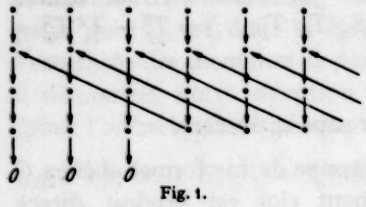


Fig. 1.

Pour démontrer notre assertion, nous remarquerons qu'en supposant le problème résolu, l'hyperalgèbre de $G_{n,m,r}$ sera l'*algèbre quotient* de l'algèbre de polynômes P par l'*idéal* $\mathfrak{a}$ engendré par les polynômes

$$(42)\qquad \begin{cases} T_h^{p^m} & \text{pour } 0 \leq h < r \\ T_h^{p^m} - T_{h-r}^{p^m} & \text{pour } h \geq r. \end{cases}$$

Posons $\overline{P} = P/\mathfrak{a}$, et désignons par $\overline{T}_h$ la classe mod. $\mathfrak{a}$ de T_h. L'algèbre $\overline{P}$ est évidemment engendrée par les éléments $X_{hi} = \overline{T}_h^{p^i}$ pour $0 \leq i \leq n-1$ et $h \geq 0$, et par suite aussi par les éléments

$$(43)\qquad \overline{E}_\alpha(\overline{T}_0, \overline{T}_1, \ldots) = \prod_{i=0}^{n-1} E_{\alpha_i}(\overline{T}_0^{p^i}, \overline{T}_1^{p^i}, \ldots, \overline{T}_h^{p^i}, \ldots)$$

où $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ ne parcourt cette fois que l'ensemble des indices «de dimension n». En fait, nous allons voir que les $\overline{E}_\alpha$ forment une *base* de $\overline{P}$ (sur le corps $\mathbf{F}_p$). Admettons ce point pour un instant; alors on peut écrire

$$(44)\qquad \overline{E}_\alpha \overline{E}_\beta = \sum_\gamma \overline{d}_{\alpha\beta\gamma} \overline{E}_\gamma \qquad (\overline{d}_{\alpha\beta\gamma} \in \mathbf{F}_p)$$

et ces identités peuvent être interprétées comme des *congruences*

$$(45)\qquad E_\alpha(T_0, T_1, \ldots)\, E_\beta(T_0, T_1, \ldots) \equiv \sum_\gamma \overline{d}_{\alpha\beta\gamma} E_\gamma(T_0, T_1, \ldots) \pmod{\mathfrak{a}}.$$

Posons alors $\overline{E}^0_\alpha(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots) = \sum_{0 \leq \beta \leq \alpha} \overline{E}_\beta(\overline{T}_0, \overline{T}_1, \ldots)\, \overline{E}_{\alpha-\beta}(\overline{U}_0, \overline{U}_1, \ldots)$.

Il est clair que pour démontrer l'existence du groupe $G_{n,m,r}$, il suffit, en vertu de la prop. 1, de prouver que les $\overline{E}^0_\alpha$ satisfont aux *mêmes* relations (44) que les $\overline{E}_\alpha$. Or, en vertu de (38), on a

$$\overline{E}^0_\alpha(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots) = E_\alpha(\overline{\Phi}_0, \overline{\Phi}_1, \ldots)$$

où on désigne par $\overline{\Phi}_h$ le polynôme $\varphi_h(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots)$. Cela étant, si, dans la congruence polynomiale (45), on substitue $\overline{\Phi}_h$ à chaque T_h, on voit que la différence

$$\overline{E}^0_\alpha \overline{E}^0_\beta - \sum_\gamma \overline{d}_{\alpha\beta\gamma} \overline{E}^0_\gamma$$

est combinaison linéaire, à coefficients dans $\overline{P}$, des éléments $\overline{\Phi}_h^{p^m}$ $(0 \leq h < r)$ et $\overline{\Phi}_h^{p^m} - \overline{\Phi}_{h-r}^{p^m}$ (pour $h \geq r$); tout revient à voir que ces différences sont *nulles*.

Mais cela revient à démontrer que l'on a

$$\varphi_h\Big(\overbrace{0,\dots,0}^{r\text{ fois}}, T_0^{p^m},\dots,T_h^{p^m},\dots;\overbrace{0,\dots,0}^{r\text{ fois}},U_0^{p^m},\dots,U_h^{p^m},\dots\Big) = \varphi_{h-r}(T_0^{p^m},\dots,T_h^{p^m},\dots;U_0^{p^m},\dots,U_h^{p^m}\dots)$$

identité qui résulte de (21) (on a naturellement utilisé le fait que tous les coefficients des polynômes qui interviennent sont dans $\boldsymbol{F}_p$).

Nous sommes donc ramenés à démontrer que les $\overline{E}_\alpha$ forment une base de $\overline{P}$. Pour cela, prouvons d'abord que les formules (40) définissent bien la table de multiplication d'une algèbre associative ayant pour base les monômes X_α. En supposant le problème résolu, on voit immédiatement (notamment sur le diagramme) que l'algèbre cherchée se décomposera en produit tensoriel de r sous-algèbres, dont chacune sera isomorphe à une algèbre engendrée par des éléments X_{hi} satisfaisant aux relations (40) où on a fait $r=1$; tout revient donc à considérer ce cas particulier. Pour cela, envisageons d'abord le cas plus particulier où $r=1$, $m=0$; $\overline{P}$ apparaît alors comme «limite inductive» de la suite d'algèbres

$$A_0 = \boldsymbol{F}_p[S_0]/(S_0^p),\ A_1 = A_0[S_1]/(S_1^p - s_0),\dots,A_k = A_{k-1}[S_k]/(S_k^p - s_{k-1}),\dots$$

où s_{k-1} désigne la classe de S_{k-1} dans A_{k-1}. Par récurrence sur k, on voit aisément que cette algèbre particulière, que nous désignerons par B, admet pour base les monômes s_α par rapport aux s_k, l'exposant de chacun des s_k étant $\leq p-1$. Notre assertion est donc établie dans ce cas. Pour $r=1$ et $0<m\leq n-1$, formons le «produit tensoriel infini» Q de B et d'une infinité d'algèbres C_k, dont chacune est isomorphe à $\boldsymbol{F}_p[T]/(T^{p^m})$, et a donc une base formée des puissances t_k^k $(0\leq k\leq p^m-1)$ d'un élément t_k avec $t_k^{p^m}=0$[2]). L'algèbre Q a donc une base formée des monômes $s_\alpha t_\beta$, où dans t_β l'exposant de chaque t_h est $\leq p^m-1$. Posons alors

$$e_0 = s_{n-1}+t_0,\quad e_1 = s_{2(n-m)+m-1}+t_1,\dots,\quad e_h = s_{(h+1)(n-m)+m-1}+t_h,\dots.$$

Comme on a $e_h^{p^m} = s_{(h+1)(n-m)-1}$, il résulte de ce qui précède que les monômes e_α, où l'exposant de chaque e_h est $\leq p^n-1$, forment une base de l'algèbre Q, et comme $e_h^{p^m} = e_{h-1}^{p^m}$ pour $h\geq 1$, $e_0^{p^n}=0$, Q est isomorphe à l'algèbre $\overline{P}=P/\mathfrak{a}$ lorsque $r=1$.

Pour établir enfin que les $\overline{E}_\alpha$ forment une base de $\overline{P}$, il suffit de remarquer que, d'après (43) et le raisonnement que nous venons de faire, $\overline{E}_\alpha$ s'exprime d'une seule façon comme polynôme par rapport aux $\overline{T}_h$, l'exposant de chacun des $\overline{T}_h$ étant toujours $\leq p^n-1$. En outre, si on ordonne encore ces monômes par rapport aux $\overline{T}_h$ suivant l'ordre lexicographique inverse, le «premier» terme de $\overline{E}_\alpha$ provient du «premier» terme de E_α: cela résulte en effet de ce

[2]) Intuitivement, l'idée de la démonstration consiste à considérer chaque «composante connexe» du diagramme; une telle composante consiste en un graphe linéaire infini (correspondant à l'algèbre B) auquel sont «attachés» aux «noeuds» espacés de $n-m$ en $n-m$, des graphes linéaires de longueur m (correspondant aux algèbres C_k).

que dans le «premier» terme de $\bar{E}_\alpha$ chaque T_h figure à une puissance $\leq p^n - 1$, puisque $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ n'a que n composantes, et du fait que les réductions qui s'opèrent dans le passage au quotient, en vertu des relations $T_h^{p^n} = T_{h-1}^{p^n}$, ne peuvent jamais qu'*augmenter* la place d'un monôme quand on passe de E_α à $\bar{E}_\alpha$. L'indépendance linéaire des $\bar{E}_\alpha$ en résulte alors comme au n° 8, et comme le nombre des $\bar{E}_\alpha$ de hauteur donnée est égal au nombre des monômes X_α de même hauteur, les $\bar{E}_\alpha$ forment une base de $\bar{P}$, ce qui achève de démontrer l'existence des groupes $G_{n,m,r}$.

On notera que, par définition du groupe $G_{n,m,r}$, l'opérateur de translation dans ce groupe est donné explicitement par la formule

$$R_y = \prod_{i=0}^{n-1} \mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^h} X_{hi}, \ldots) \tag{46}$$

où bien entendu il faut réduire les puissances des X_{hi} à des exposants $< p$ compte tenu des formules (40). En outre, si u' est l'homomorphisme canonique de P sur $\bar{P} = P/\mathfrak{a}$, u' transforme la sous-algèbre $\mathfrak{U}^0$ du produit tensoriel $\mathfrak{U} \otimes \mathfrak{U}$ ($\mathfrak{U}$ hyperalgèbre de U) en la sous-algèbre du produit tensoriel $\mathfrak{G}_{n,m,r} \otimes \mathfrak{G}_{n,m,r}$ ($\mathfrak{G}_{n,m,r}$ hyperalgèbre de $G_{n,m,r}$) engendrée par les $\bar{\Phi}_h$ [en raison de (38)]. Ce qui précède prouve donc (prop. 2) que u' est le dérivé d'un *homomorphisme* u du groupe U sur $G_{n,m,r}$.

Lorsque $n = 1$, on a nécessairement $m = 0$ et les formules (40) montrent que le groupe $G_{1,0,r}$ est *isomorphe au groupe* I_r défini dans [5, n° 8] et dont l'existence se trouve ainsi établie.

On notera d'autre part que $G_{n,m,r}$ et $G_{n',m',r'}$ *ne peuvent être isomorphes que si* $n = n'$, $m = m'$ *et* $r = r'$. En effet, l'*indice de nilpotence* de l'algèbre $\mathfrak{s}_h$ des semi-dérivations spéciales invariantes de hauteur $< h$ (plus petit exposant e tel que $Y^e = 0$ pour tout élément Y de l'algèbre) est égal à p^n pour $h \leq r$, à p^{2n-m} pour $h = r+1$, et par suite les nombres m et r sont déterminés intrinsèquement par le groupe $G_{n,m,r}$.

10.

Comme deuxième exemple de groupes obtenus à partir du groupe U, considérons des groupes de dimension 2, dont l'algèbre de Lie $\mathfrak{g}_0$ est «complètement triviale», c'est-à-dire telle que la puissance p-ème de tout élément de $\mathfrak{g}_0$ soit nulle. On a évidemment, parmi ces groupes, tous les produits directs $I_r \times I_s$ ($r > 0$, $s > 0$), $I_r \times W_1$ et $W_1 \times W_1$. En outre, nous allons voir que l'on a encore des groupes correspondant aux trois types de diagrammes suivants:

I. *Type «strie»*:

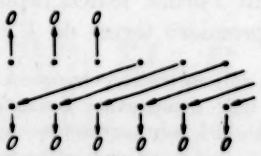


Fig. 2.

En formules, cela correspond aux relations

$$(47)\qquad \begin{cases} X^p_{h0}=0 \quad \text{pour} \quad h<r, \quad X^p_{h+r,0}=X_{h1} \quad \text{pour} \quad h\geq 0 \\ X^p_{h1}=0 \quad \text{pour tout} \quad h\geq 0. \end{cases}$$

II. *Type «arbre»*:

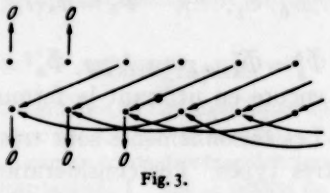


Fig. 3.

ou, en formules:

$$(48)\qquad \begin{cases} X^p_{h0}=0 \quad \text{pour} \quad h<r, \quad X^p_{h+r,0}=X_{h1} \quad \text{pour} \quad h\geq 0 \\ X^p_{h1}=0 \quad \text{pour} \quad h<s, \quad X^p_{h+s,1}=X_{h1} \quad \text{pour} \quad h\geq 0. \end{cases}$$

III. *Type «tresse»*:

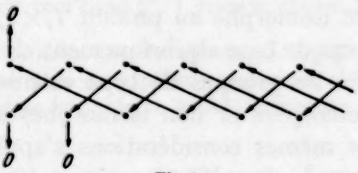


Fig. 4.

ou, en formules:

$$(49)\qquad \begin{cases} X^p_{h0}=0 \quad \text{pour} \quad h<r, \quad X^p_{h+r,0}=X_{h1} \quad \text{pour} \quad h\geq 0 \\ X^p_{h1}=0 \quad \text{pour} \quad h<s, \quad X^p_{h+s,1}=X_{h0} \quad \text{pour} \quad h\geq 0. \end{cases}$$

L'existence des groupes de ces types se démontre en partant cette fois du groupe produit $U\times U$, dont l'hyperalgèbre est $P\otimes P$, et peut par suite être considérée comme algèbre des polynômes à 2 séries de variables $T_0, T_1, \ldots, T'_0, T'_1, \ldots$, ayant pour base les polynômes $E_\alpha(T_0, T_1, \ldots)E_\beta(T'_0, T'_1, \ldots)$. Si on veut par exemple démontrer l'existence des groupes du type «tresse», on considérera l'idéal $\mathfrak{b}$ de $P\otimes P$ engendré par les polynômes

$$T^p_h \quad \text{pour} \quad h<r, \qquad T^p_h - T'_{h-r} \quad \text{pour} \quad h\geq r$$

$$T'^p_h \quad \text{pour} \quad h<s, \qquad T'^p_h - T_{h-s} \quad \text{pour} \quad h\geq s.$$

Si $\overline{T}_h, \overline{T}'_h$ sont les classes de T_h, T'_h mod. $\mathfrak{b}$, on montre d'abord que les éléments $\overline{E}_\lambda(\overline{T}_0, \overline{T}_1, \ldots; \overline{T}'_0, \overline{T}'_1, \ldots) = E_k(\overline{T}_0, \overline{T}_1, \ldots)E_l(\overline{T}'_0, \overline{T}'_1, \ldots)$, où $\lambda=(k,l)$ parcourt tous les couples d'entiers ≥ 0, forment une *base* de $(P\otimes P)/\mathfrak{b}$; la méthode est analogue à celle du n° 9, et consiste à «dérouler» la «tresse» en ses composantes connexes, autrement dit, à prouver que $(P\otimes P)/\mathfrak{b}$ est isomorphe au produit tensoriel de $r+s$ algèbres identiques à l'algèbre B définie au n° 9. Ceci montre que les monômes X_α par rapport aux $X_{h0}=\overline{T}_h$ et $X_{h1}=\overline{T}'_h$, ayant tous leurs exposants $<p$, forment une base d'une algèbre isomorphe à $(P\otimes P)/\mathfrak{b}$, et vérifiant (49). Pour voir ensuite que les $\overline{E}_\lambda$ forment aussi une base de cette algèbre, il suffit d'ordonner les $\overline{T}_h$ et $\overline{T}'_h$ suivant l'ordre: $\overline{T}_0, \overline{T}'_0, \overline{T}_1, \overline{T}'_1, \ldots, \overline{T}_h, \overline{T}'_h, \ldots$, puis les monômes X_α suivant l'ordre lexicographique inverse (ce n'est autre que l'ordre lexicographique inverse «par

étages» décrit au n° 3); le raisonnement se poursuit alors comme au n° 9. Reste enfin à montrer que la prop. 1 est applicable; on voit comme au n° 9 que cela revient à prouver que, si on pose

$$\overline{\Phi}_h = \varphi_h(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots), \quad \overline{\Phi}'_h = \varphi_h(\overline{T}'_0, \overline{T}'_1, \ldots; \overline{U}'_0, \overline{U}'_1, \ldots)$$

on a $\overline{\Phi}_h^p = 0$ pour $h < r$, $\overline{\Phi}_h^p = \overline{\Phi}'_{h-r}$ pour $h \geq r$, $\overline{\Phi}_h'^p = 0$ pour $h < s$, $\overline{\Phi}_h'^p = \overline{\Phi}_{h-s}$ pour $h \geq s$; cela se fait encore en utilisant la formule (21).

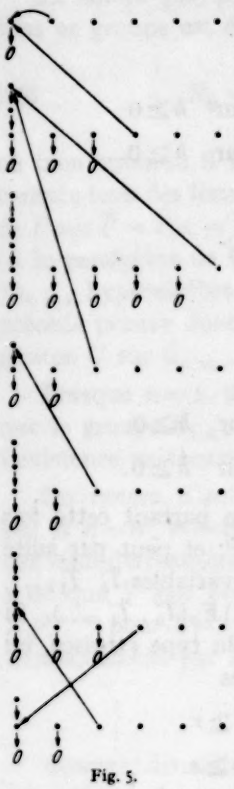


Fig. 5.

Les raisonnements sont très analogues pour les deux autres types. Les considérations d'indice de nilpotence montrent en outre que les groupes de type «tresse dissymétrique» ($r < s$) ne sont pas isomorphes deux à deux ni à aucun groupe des autres types; par contre, on peut prouver que le type «tresse symétrique» ($r = s$) est isomorphe au produit $I_r \times I_r$ (tout au moins sur un corps de base algébriquement clos). De même, on montre que les groupes de type «strie» sont deux à deux non isomorphes et non isomorphes aux autres types; enfin, les mêmes considérations s'appliquent au type «arbre» lorsque $r < s$; j'ignore par contre si pour $r \geq s$, un groupe du type «arbre» ne peut être isomorphe à $W_1 \times I_s$.

11.

Les considérations des deux n^{os} précédents admettent évidemment une généralisation étendue. On forme ainsi des groupes dont l'algèbre de Lie $\mathfrak{g}_0$ est somme directe d'un nombre arbitraire d'algèbres *indécomposables* (on sait [5, n° 11] que toute algèbre de Lie «sans noyau» est de cette forme). A chacune de ces algèbres de Lie indécomposables, nous attacherons un «diagramme simplifié» ne comportant qu'une ligne horizontale et une ligne de flèches verticales; puis nous relierons ces diagrammes par des flèches obliques (éventuellement horizontales), suivant les pricipes suivants: 1° de chaque ligne horizontale, il ne part qu'une seule flèche oblique au plus; 2° la flèche oblique aboutit en un point d'une ligne verticale. On complète ensuite le diagramme en opérant sur chaque flèche une infinité de translations d'une unité vers la droite (non représentées sur la figure qui suit) et en ajoutant à chaque ligne horizontale des flèches aboutissant à 0, avant le début de la première flèche oblique:

L'interprétation des diagrammes est la même que ci-dessus: s'il y a k diagrammes, dont les nombres de lignes sont n_i ($1 \leq i \leq k$), les éléments X_{hij} ($1 \leq i \leq k$, $1 \leq j \leq n_i$) des h-èmes colonnes de ces diagrammes sont les semi-dérivations invariantes de hauteur h telles que $X_{hij}(e) = D_{hij}(e)$, et une flèche joint une semi-dérivation à sa puissance p-ème.

Il est naturel de conjecturer que l'on obtient ainsi *tous* les groupes abéliens radicaux (à des isomorphismes près); mais je ne sais pour le moment confirmer ou infirmer cette conjecture[3]). Au cas où elle serait exacte, une étude détaillée des diagrammes précédents et des isomorphismes possibles entre les groupes qui leur correspondent, serait naturellement indispensable.

Bibliographie.

[*1*] DIEUDONNÉ, J.: Sur les groupes de Lie algébriques sur un corps de caractéristique $p > 0$. Rend. Circ. Mat. Palermo (2) **1**, 380—402 (1952). — [*2*] DIEUDONNÉ, J.: Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$. Comm. Math. Helv. **28**, 87—118 (1954). — [*3*] DIEUDONNÉ, J.: Sur la notion de variables canoniques. Anais Acad. brasil. Ciencias **27** (1955). — [*4*] DIEUDONNÉ, J.: Le calcul différentiel dans les corps de caractéristique $p > 0$. Proc. Int. Congress of Math., Amsterdam **1** (1955). — [*5*] DIEUDONNÉ, J.: Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (II). Amer. J. Math. **77**, 218—244 (1955). — [*6*] WITT, E.: Zyklische Körper und Algebren der Charakteristik p vom Grad p^n. J. reine u. angew. Math. **176**, 126—140 (1937).

[3]) *Note ajoutée pendant la correction des épreuves.* Depuis que cet article a été écrit, j'ai obtenu un procédé général de description des groupes abéliens qui comprend comme cas particuliers ceux envisagés ici; voir un article à paraître dans l'*American Journal of Mathematics*.

Dept. of Math., Northwestern University, Evanston, Ill. (USA).

(Eingegangen am 8. November 1954.)

NEUMANN, B. H.
Math. Zeitschr. Bd. 63, S. 76—96 (1955)

Groups with finite classes of conjugate subgroups.

(In memoriam ISSAI SCHUR.)

By

B. H. NEUMANN.

§ 1. Introduction.

In the present paper we shall be concerned with the interrelations of various classes of groups defined by certain properties, and to avoid repetitive circumlocutions we make the following convention. The properties of groups we are mainly interested in will be denoted by short sequences of capitals. If PR is a property of groups, the groups that have the property PR will be called PR-groups, and the class of all PR-groups will be denoted by $[PR]$. Thus, for instance, we denote by FG the property to be "finitely generated" (more accurately: to be capable of being finitely generated; for such groups can, of course, be given in terms of an infinite set of generators). Then $[FG]$ is the class of all FG-groups, that is the class of all finitely generated groups.

The most important properties of groups G to be considered are the following[1]).

FC: The classes of conjugate elements of G are finite.

FD: The derived group G' is finite.

FIZ: The centre has finite index in G.

X: The classes of conjugate subgroups of G are finite.

Y: Every subgroup H of G has finite index in some normal subgroup of G (and therefore, in particular, in its normal closure H^G in G).

FC-groups have recently received attention from several authors[2]). FD is a more restrictive property than FC, that is to say, all FD-groups are FC-groups; FIZ is still more restrictive, that is to say, all FIZ-groups are FD-groups. For finitely generated groups the three properties are co-extensive. We can formulate these (known) relations thus:

$$[FC] > [FD] > [FIZ]; \tag{1.1}$$

$$[FG] \cap [FC] = [FG] \cap [FD] = [FG] \cap [FIZ]. \tag{1.2}$$

The relations of X and Y to each other and to the other properties are far from obvious. One might try to be guided by the extreme case that arises if "finite" is replaced by "of cardinal 1". Then FC-groups, FD-groups and FIZ-groups all simply become abelian groups, whereas X-groups and Y-

[1]) Further properties, namely BFC, BX, BY, are defined at the end of this section. The classes $[H]$—$[K]$ introduced in §§ 5—10 are of ephemeral significance only.

[2]) Cf. BAER [1], NEUMANN [9], HAIMO [5], [6], ERDÖS [4] and the literature there quoted; also NEUMANN [10].

groups become the groups in which every subgroup is normal. These groups, first studied by DEDEKIND [3], form a somewhat larger class than that of abelian groups, comprising the abelian and the hamiltonian groups. It is then not unnatural to expect that analogously $[FC]$ is contained in $[X]$ and in $[Y]$. This is, however, not the case; on the contrary, we shall show that the X-groups are precisely the FIZ-groups, and the Y-groups are precisely the FD-groups. Thus the interrelations of these classes of groups can be completely—and simply—described. In particular it follows that every X-group is a Y-group; I know no obvious reason why this should be so, nor do I know any direct proof.

A corollary of our results is that in an X-group the classes of conjugate subgroups are not only finite, but boundedly (or uniformly) finite; and that in a Y-group every subgroup has not only finite, but boundedly finite index in some normal subgroup. Let us denote the properties obtained from FC, X, and Y by strengthening "finite" to "boundedly finite" by BFC, BX, BY respectively. Then $[BFC]$ is a proper subclass of $[FC]$; in fact it has been shown [*10*, Theorem 3.1] that $[BFC]=[FD]$, and we shall here rederive this result by a quite different method. On the other hand we show that $[BX]=[X]$ and $[BY]=[Y]$. The results of this paper can then be summarized in the form

$$(1.3) \qquad [FC] > [BFC] = [FD] = [Y] = [BY] > [FIZ] = [X] = [BX].$$

§ 2. Preliminaries.

We use the following notation. If G is a group, then G' is its derived group, Z its centre. If S is a subset of G, then $|S|$ denotes its cardinal; $\{S\}$ is the subgroup it generates; $N(S)$ is its normalizer and $C(S)$ its centralizer in G. If S consists of $g_1, g_2, \ldots, g_n$, then we also write $\{g_1, g_2, \ldots, g_n\}$ for $\{S\}$, $N(g_1, g_2, \ldots, g_n)$ for $N(S)$, and $C(g_1, g_2, \ldots, g_n)$ for $C(S)$. We note that

$$C(\{S\}) = C(S) = \bigcap_{s \in S} C(s),$$

and if S is a generating set of G, then

$$C(S) = C(G) = Z.$$

In general, however, $N(\{S\})$ is not the same as $N(S)$. For a single element, $N(g) = C(g)$.

If H is a subgroup of G, then $|G:H|$ denotes its index in G, and H^G its normal closure in G, that is the least normal subgroup of G containing H. The unit element of all the groups that occur will be denoted by 1.

The class of conjugates of a subset S of G consists of the subsets $g^{-1}Sg$, where g ranges over G. Its cardinal is the index of the normalizer, $|G:N(S)|$. In particular the cardinal of the class of conjugates of an element g is $|G:C(g)|$.

The following known facts on FC-groups will be required [*9*, § 5].

Lemma 2.1. *If G is an FC-group then the elements of finite order form a subgroup P of G; if G is finitely generated then P is finite; P contains G', and*

thus all commutators in G have finite order. Every finite set of elements of finite order generates a finite subgroup, and is even contained in a finite normal subgroup. If S is a finite subset of G then $|G:C(S)|$ is finite.

Lemma 2.2. *Each of the classes $[FC]$, $[BFC]$, $[FD]$, $[FIZ]$, $[X]$, $[BX]$, $[Y]$, $[BY]$ is closed under the operations of forming subgroups and homomorphic images.*

The proof is obvious and omitted.

Lemma 2.3. *Every FIZ-group is an X-group and even a BX-group; every FD-group is a Y-group and even a BY-group; that is to say,*

$$[FIZ] \subseteqq [BX] \subseteqq [X]; \quad [FD] \subseteqq [BY] \subseteqq [Y].$$

Proof. The centre Z is contained in every centralizer and *a fortiori* in every normalizer; thus if $G \in [FIZ]$, that is if $|G:Z|$ is finite, then so is $|G:N(H)|$ for every subgroup H, and $G \in [X]$; in fact $|G:Z|$ is a bound for the number of conjugates of every subgroup, and thus $G \in [BX]$. Hence $[FIZ] \subseteqq [BX]$. Next, if H is a subgroup of G, then HG' is a normal subgroup of G. Also $|HG':H| \leq |G'|$; if $G \in [FD]$, that is if $|G'|$ is finite, then so is $|HG':H|$ for every H, and $G \in [Y]$; in fact $|G'|$ is a bound for the index of H in one normal subgroup of G, and thus $G \in [BY]$. Thus $[FD] \subseteqq [BY]$, and the lemma follows.

§ 3. X-groups and Y-groups are FC-groups.

Lemma 3.1. *If g is an element of a group G then $|N(\{g\}):C(g)|$ is finite.*

Proof. The cyclic group $\{g\}$ is normal in $N(\{g\})$. Each conjugate of g in $N(\{g\})$ then lies in $\{g\}$ and generates $\{g\}$; but a cyclic group has only a finite number of distinct single generators [namely $\varphi(n)$ if $\{g\}$ is of finite order n, and 2 if $\{g\}$ is of infinite order]. Hence g has only a finite number of distinct conjugates in $N(\{g\})$. Finally we observe that the centralizer of g in $N(\{g\})$ is the same as that in G; and the lemma follows.

Lemma 3.2. *Every X-group is an FC-group.*

Proof. If $G \in [X]$ and $g \in G$, then $|G:N(\{g\})|$ is finite because $\{g\}$ has only finitely many distinct conjugates, and $|N(\{g\}):C(g)|$ is finite by the preceding lemma. Hence

$$|G:C(g)| = |G:N(\{g\})| \cdot |N(\{g\}):C(g)|$$

is finite, and g has only a finite number of distinct conjugates. This is true for every $g \in G$; hence $G \in [FC]$, and the lemma follows.

Corollary 3.3. $$[FG] \cap [X] = [FG] \cap [FC].$$

This follows from (1.2) and Lemmas 2.3 and 3.2.

Lemma 3.4[3]). *A finitely generated group H contains only finitely many subgroups K of given finite index $|H:K| = k$.*

[3]) This lemma is known; cf. MARSHALL HALL [7], who determines the precise number when H is free, and thereby provides a bound for arbitrary H. Cf. also BAER [2, p. 331].

Proof. Choose a fixed finite set of generators of H. Every subgroup K of index k determines an equivalence relation on H, namely the right congruence modulo K. This equivalence is completely determined when it is specified for all words of length not exceeding some fixed bound l, and every such word of length l is equivalent to a shorter word. If $h_1 h_2 \dots h_k$ is a word of length k (the h_i being generators or their inverses), then among the $k+1$ elements

$$1, h_1, h_1 h_2, \dots, h_1 h_2 \dots h_k$$

there must be two that are right congruent modulo K, and it easily follows that then $h_1 h_2 \dots h_k$ is right congruent to a shorter word modulo K. Hence we can put $l = k$. Conversely, the right congruence determines K, because K simply consists of all elements right congruent to 1. As there are only a finite number of elements of length not exceeding k, and only a finite number of equivalence relations on a finite set, the number of subgroups K is finite, and the lemma follows.

Lemma 3.5. *If a finitely generated subgroup K of G has finite index in a normal subgroup H of G, then K has only finitely many distinct conjugates in G.*

Proof[4]). We first note that H is itself finitely generated, namely by a finite generating set of K and one representative out of each right coset of K in H. Now every conjugate of K in G is a subgroup of H and has the same index $k = |H:K|$ in H; by the preceding lemma there are only a finite number of such subgroups, and the lemma follows.

Lemma 3.6. *If a cyclic group $\{g\}$ has finite index in its normal closure $\{g\}^G$, then its generator g has only finitely many distinct conjugates in G.*

Proof. By the preceding lemma, $|G:N(\{g\})|$ is finite, and then $|G:C(g)|$ is finite by Lemma 3.1.

Lemma 3.7. *Every Y-group is an FC-group.*

This is an immediate consequence of the definition of $[Y]$ together with Lemma 3.6.

Corollary 3.8. $$[FG] \cap [Y] = [FG] \cap [FC].$$

This follows from (1.2) and Lemmas 2.3 and 3.7.

§ 4. Some properties of FC-groups.

In the present section we show that if an FC-group G contains a sufficiently "large" subgroup in $[FD]$ or in $[FIZ]$, then G is itself an FD-group or FIZ-group, respectively. We also demonstrate a method of replacing a group by a factor group without disturbing any of the properties we are interested in. All these results will be found useful at a later stage.

Lemma 4.1. *Let $G \in [FC]$ possess a subgroup $H \in [FD]$ and a finite subset S such that H and S together generate G. Then $G \in [FD]$.*

[4]) I am indebted to GRAHAM HIGMAN for a simplification of the proof of this lemma.

Proof. The derived group G' is the normal closure of the group generated by the commutators $[h, h']$, $[h, s]$, and $[s, s']$, where h, h' range over H and s, s' over S. There are only finitely many $[h, h']$ because H' is finite; there are only finitely many commutators $[g, s]$ for any one s and for g ranging over G, because such a commutator is the quotient of a conjugate of s—capable of only finitely many different values—and s; thus there are altogether only finitely many commutators of the form $[g, s]$ with s ranging over S. It follows that G' is the normal closure of a finite set of elements of finite order and hence of finite order itself (cf. Lemma 2.1). Thus $G \in [FD]$ and the lemma follows.

Corollary 4.2. *If G is an FC-group but not an FD-group, then no subgroup of finite index in G is an FD-group.*

Lemma 4.3. *Let $G \in [FC]$ possess a subgroup $H \in [FIZ]$ and a finite subset S such that H and S together generate G. Then $G \in [FIZ]$.*

Proof. Denote by K the intersection of H and the centre Z of G:

$$K = H \cap Z.$$

As $G = \{H, S\}$, then

$$Z = C(H) \cap C(S),$$

and therefore

$$K = H \cap C(H) \cap C(S) = (H \cap C(H)) \cap (H \cap C(S)).$$

Now $H \cap C(H)$ is the centre of H, and this has finite index in H by the assumption that $H \in [FIZ]$. Also

$$|H : H \cap C(S)| \leq |G : C(S)|,$$

and this is finite by Lemma 2.1, as S is a finite set. Hence K, as the intersection of two subgroups of finite index in H, has finite index in H. Denote by T a subset of H which together with K generates H. Then $G = \{K, T, S\}$ and

$$Z = C(K) \cap C(T) \cap C(S).$$

Here $C(S)$ and $C(T)$ have finite index in G because S and T are finite sets (cf. Lemma 2.1); also $C(K) = G$, as $K \subseteq Z$. It follows that $|G : Z|$ is finite, and $G \in [FIZ]$, as asserted.

Corollary 4.4. *If G is an FC-group but not an FIZ-group, then no subgroup of finite index in G is an FIZ-group.*

Lemma 4.5. *In an FC-group every element has finite order modulo the centre.*

Proof. Let $g \in G \in [FC]$; then $|G : C(g)|$ is finite. Let $g_1, g_2, \ldots, g_n$ be a complete set of coset representatives of $C(g)$ in G. Then also $|G : C(g_1, g_2, \ldots, g_n)|$ is finite. Hence there is a positive integer m such that $g^m \in C(g_1, g_2, \ldots, g_n)$. As g^m commutes with all the elements that commute with g, that is with all the elements of $C(g)$, and also with $g_1, g_2, \ldots, g_n$, then $g^m \in Z$, and the lemma follows.

Definition. Let φ be a homomorphic mapping of a group G onto a group $G^0 = G\varphi$, and let K be the kernel of φ. If $K \cap G' = \{1\}$ and hence $K \subseteq Z$, then we call φ an *isoclinic* homomorphism.

If φ is an isoclinic homomorphism, then the centre Z^0 of G^0 is $Z\varphi$; for if $g \in G$ is not in the centre Z, then there is an element $h \in G$ such that $[g, h] \neq 1$. But then $[g, h] \notin K$, because $K \cap G' = \{1\}$, and therefore $[g\varphi, h\varphi] = [g, h]\varphi \neq 1$. Thus $g\varphi$ is outside the centre of $G\varphi$. It follows that the factor groups of the centres of G and G^0 are isomorphic; for

$$G/Z \cong G\varphi/Z\varphi = G^0/Z^0$$

as Z contains the kernel of φ. In fact φ itself induces an isomorphism of these factor groups; and this has clearly the property of inducing an isomorphism of the derived groups G' and $G^{0\prime}$ also, because $K \cap G' = \{1\}$ entails

$$G^{0\prime} = (G\varphi)' = G'\varphi.$$

All this together shows that G and G^0 are isoclinic in the sense of PHILIP HALL [8]. The isoclinic homomorphism is not the most general isoclinism between groups, but it is sufficient for our purposes.

Lemma 4.6. *Let φ be an isoclinic homomorphism of a group G onto $G^0 = G\varphi$. Then if $G^0 \in [PR]$ where $[PR]$ is one of the classes $[FC]$, $[FD]$, $[FIZ]$, then also $G \in [PR]$.*

Proof. As the derived groups of G^0 and of G are isomorphic, $G^0 \in [FD]$ implies $G \in [FD]$. As the factor groups of the centres of G^0 and G are isomorphic, $G^0 \in [FIZ]$ implies $G \in [FIZ]$. We next show that $g \in G$ has precisely as many distinct conjugates as $g^0 = g\varphi$ has in G^0. Let $h \notin C(g)$. Then $[h, g] \neq 1$, and therefore $[h, g] \notin K$; thus $[h\varphi, g\varphi] = [h, g]\varphi \neq 1$, and $h^0 = h\varphi$ is not in the centralizer $C^0(g^0)$ of g^0 in G^0. Thus nothing outside $C(g)\varphi$ can belong to $C^0(g^0)$; but clearly $C(g)\varphi \subseteq C^0(g^0)$. As, moreover, $C(g)$ contains Z, and thus also K, then

$$|G : C(g)| = |G\varphi : C(g)\varphi| = |G^0 : C^0(g^0)|.$$

It follows that the class of conjugates of an element $g \in G$ has the same cardinal as that of g^0 in G^0; in particular, if $G^0 \in [FC]$, then $G \in [FC]$. This completes the proof of the lemma. It may be noted that the lemma provides a partial converse of Lemma 2.2.

Lemma 4.7. *To every group G there is an isoclinic homomorphism onto a group G^0 in which every element $z^0 \neq 1$ of the centre Z^0 has a power $z^{0m} \neq 1$ in the derived group $G^{0\prime}$.*

Proof. Among the kernels of isoclinic homomorphisms of G, that is among the subgroups of Z which have trivial intersection with G', we choose a maximal one and denote it by K. The existence of such a maximal subgroup follows by a familiar application of ZORN's Lemma. We put $G^0 = G/K$; the canonic homomorphism of G onto G^0 is then isoclinic. If Z^0 is the centre of G^0, then we know (from a remark following the definition of an isoclinic homomorphism) that $Z^0 = Z/K$. Now every non-trivial subgroup of Z^0 must intersect the derived group $G^{0\prime}$ non-trivially; this is an easy consequence of the maximality of K. Thus if $z^0 \in Z^0$, $z^0 \neq 1$, then there is an integer m such that $1 \neq z^{0m} \in G^{0\prime}$, and the lemma follows.

Corollary 4.8. *If $G \in [FC]$, then G^0—chosen according to Lemma 4.7—is periodic, that is to say, all its elements have finite order; and G^0 is even locally finite, that is to say, every finite subset is contained in a finite subgroup.*

Proof. As G^0 is a homomorphic map of G, then $G^0 \in [FC]$. By Lemma 4.5 every element $g^0 \in G^0$ has finite order modulo Z^0. If $g^{0m} = z^0 \in Z^0$, then g^0 has certainly finite order if $z^0 = 1$; if $z^0 \neq 1$ then $g^{0mm} = z^{0m} = c^0 \in G^{0\prime}$ by Lemma 4.7, and c^0 has finite order by Lemma 2.1. Hence g^0 has finite order, and by a further application of Lemma 2.1 the corollary follows.

§ 5. The class *[H]*.

In this and the following four sections we introduce and study some rather restricted classes of FC-groups which will later be shown to be typical of certain subgroups, or certain factor groups of subgroups, that occur among the groups we are concerned with.

We denote by $[H]$ the class of groups G generated by (denumerably) infinitely many elements

$$a_1, b_1, c_1, a_2, b_2, c_2, \ldots$$

subject to the relations

$$(5.1) \qquad [a_i, b_i] = c_i \neq 1, \quad [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1,$$

where $i, j = 1, 2, \ldots$ and $i \neq j$; moreover each subgroup

$$N_i = \{a_i, b_i\} \,(= \{a_i, b_i, c_i\})$$

is to be an FC-group.

The relations (5.1) are not to be understood to form a system of defining relations of G: there will in general be further relations, not consequences of (5.1); but they must not imply $c_i = 1$ for any i. The sequence $a_1, b_1, c_1, a_2, b_2, c_2, \ldots$ will be called a system of H-generators of G.

The following properties of an H-group G are readily verified from the definition. The subgroups N_i are normal in G, and they jointly generate G. The centralizer $C(N_i)$ is generated by all the subgroups N_j with $j \neq i$ together with the centre of N_i. If an element of N_i is also generated by the N_j with $j \neq i$, then it belongs to the centre Z of G; for *qua* element of N_i it is permutable with the elements of all N_j other than N_i, and *qua* product of elements out of the N_j other than N_i it is also permutable with the elements of N_i. The centre Z is generated by the centres of all the N_i. We can consider G as the restricted direct product of the N_i with certain subgroups of their centres amalgamated: but this interpretation will not be required in what follows. Putting

$$F_i = N_i Z/Z,$$

one easily sees that G/Z is the restricted direct product of the groups F_i. Each F_i is a finite group, because N_i, being a finitely generated FC-group, is a FIZ-group [cf. (1.2)], and thus also $|N_i Z : Z|$ is finite. On the other hand G is not a FIZ-group; for none of $a_1, a_2, \ldots$ belong to the centre, and they

are in fact all incongruent modulo Z: if $i \neq j$, then

$$b_i^{-1} a_i a_j^{-1} b_i = a_i a_j^{-1} c_i \neq a_i a_j^{-1},$$

and thus $Z a_i \neq Z a_j$. However, G is an FC-group; for every element of G is of the form

$$g = n_1 n_2 \dots n_q, \qquad n_i \in N_i,$$

q being taken sufficiently large; the conjugates of g then are of the form

$$g' = n_1' n_2' \dots n_q'$$

where n_i' is conjugate to n_i in N_i, and thus capable of finitely many different values only. Hence the class of conjugates of g is finite, and G is an FC-group. We have thus shown

(5.2) $$[H] \leqq [FC] - [FIZ].$$

§ 6. The class $[I]$.

Next we introduce a subclass of $[H]$: we denote by $[I]$ the class of H-groups G that satisfy, in addition to (5.1), the relations

(6.1) $$[a_i, c_i] = [b_i, c_i] = 1, \qquad c_i^{p^{\varrho(i)}} = 1,$$

where $i = 1, 2, \dots$ and p is a prime number. Again there may be further relations, but they must not imply $c_i = 1$ for any i; and we may take $p^{\varrho(i)}$ to be precisely the order of c_i.

The following facts are easy consequences of the definition. In an I-group G the subgroup

$$C = \{c_1, c_2, \dots\}$$

is the derived group; it is contained in the centre, in other words G is nilpotent of class 2. If λ, μ are any integers then

$$[a_i^\lambda, b_i^\mu] = c_i^{\lambda\mu}.$$

Hence a_i (and also b_i) has order $p^{\varrho(i)}$ modulo the centre. With the notation

$$F_i = N_i Z / Z$$

of § 5, each F_i is a direct product

$$F_i = \{Z a_i\} \times \{Z b_i\}$$

of two cycles of order $p^{\varrho(i)}$; hence G/Z is the restricted direct product of cyclic groups.

We denote by A the subgroup of G generated by all a_i and by K the group generated by A and C:

$$A = \{a_1, a_2, \dots\}, \qquad K = \{A, C\} = AC.$$

These groups are abelian.

Lemma 6.2. *If*

$$a_1^{\lambda(1)} a_2^{\lambda(2)} \dots a_q^{\lambda(q)} = c \in C,$$

then $\lambda(i) \equiv 0 \bmod p^{\varrho(i)}$ $(i = 1, \dots, q)$.

Proof. Transforming by b_i we get

$$b_i^{-1} c\, b_i = c \cdot c_i^{\lambda(i)}.$$

But $c \in C \subseteq Z$; hence $c^{\lambda(i)} = 1$, and $\lambda(i)$ is divisible by the order $p^{\varrho(i)}$ of c_i.

Lemma 6.3. *Put*

$$L = \{a_1, \ldots, a_s, a_{s+1}^p, a_{s+2}^p, \ldots, a_{s+t}^p, a_{s+t+1}, a_{s+t+2}, \ldots, c_1, c_2, \ldots\};$$

let R be a subgroup of L and $Q = \{R, a_{s+1}, a_{s+2}, \ldots, a_{s+t}\}$. Then Q/R can not be generated by fewer than t elements.

Proof. It clearly suffices to prove this in the special case that $R = L$, when $Q = K$; for in any case K/L is isomorphic to a factor group of Q/R. Now K/L is an elementary abelian group of order p^{t-s}; for if

$$(L\, a_{s+1})^{\lambda(1)} (L\, a_{s+2})^{\lambda(2)} \ldots (L\, a_{s+t})^{\lambda(t)} = L$$

is a relation between the $t-s$ obvious generators of K/L, then by an easy application of Lemma 6.2

$$\lambda(1) \equiv \lambda(2) \equiv \cdots \equiv \lambda(t) \equiv 0 \pmod{p},$$

that is to say, the relation is trivial; and the lemma follows.

Lemma 6.4. *Let R be a subgroup of $\{a_1, a_2, \ldots, a_s\}$, and let $k_1, k_2, \ldots, k_u$ be u elements of order p in K but not in R. Let t be such that*

$$u < \frac{p^t - 1}{p - 1}$$

(thus $t = [\log_2 u] + 2$ will do for all p). Then there are numbers $\lambda(1), \lambda(2), \ldots, \lambda(t)$, not all divisible by p, such that the group $\{R, a^*\}$ obtained from R by adjoining the element

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)}$$

also contains none of $k_1, k_2, \ldots, k_u$.

Proof. We put $Q = \{R, a_{s+1}, a_{s+2}, \ldots, a_{s+t}\}$, and consider maximal cyclic subgroups of Q/R. If $\{R\, a^*\}$ is such a maximal cyclic subgroup, and if

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)},$$

then $\lambda(1), \lambda(2), \ldots, \lambda(t)$ can not all be divisible by p; for if $\lambda(i) = p\,\mu(i)$, and if

$$a^{**} = a_{s+1}^{\mu(1)} a_{s+2}^{\mu(2)} \ldots a_{s+t}^{\mu(1)},$$

then $\{Ra^{**}\}$ contains $\{Ra^*\}$ properly; this follows from the fact that the order of a^{**} modulo the centre Z of G is a power of p. If there is a maximal cyclic subgroup $\{Ra^*\}$ of Q/R which contains none of $Rk_1, Rk_2, \ldots, Rk_u$, then $\{R, a^*\}$ does not contain any of $k_1, k_2, \ldots, k_u$, and the lemma follows. It remains to show that such a maximal cyclic subgroup exists. If Q/R is infinite, then—being finitely generated—it contains infinite maximal cyclic subgroups, and any one of these will do, as it can not contain any element of order p. Assume then that Q/R is finite. It can evidently be generated by t elements and—by Lemma 6.3—by no fewer than t elements. Hence it is the direct product of t cyclic subgroups; these all have order divisible by p.

The elements of order p in Q/R therefore form a subgroup of order p^t; this consists of

$$\frac{p^t-1}{p-1} > u$$

cycles of order p. Thus there is at least one such cycle containing none of $Rk_1, Rk_2, \ldots, Rk_u$; and a maximal cyclic subgroup of Q/R containing such a cyclic subgroup of order p then has the required property. This completes the proof of the lemma.

Corollary 6.5. *In the lemma, the assumption on the order of $k_1, k_2, \ldots, k_u$ can be relaxed to require them only to be elements of finite orders all divisible by p.*

For then certain powers of them have order p, and a subgroup not containing these powers can not contain the elements themselves, either.

§ 7. The class [*II*].

We denote by $[II]$ the class of those H-groups G that satisfy, in addition to (5.1), the conditions

$$c_1 = c_2 = \ldots = c, \text{ say}, \qquad c^p = 1. \tag{7.1}$$

In such a group the commutator c lies in the intersection of several (in fact all) N_i and, therefore, in the centre; it follows that the conditions (6.1) are also satisfied, with $\varrho(1) = \varrho(2) = \cdots = 1$, and

$$[II] \subseteqq [I].$$

The derived group of a II-group is the cycle $\{c\}$ of order p. Hence the group is an FD-group, and using (5.2) we see that

$$[II] \subseteqq [FD] - [FIZ]. \tag{7.2}$$

The existence of II-groups is easily verified; examples are the restricted direct product of infinitely many quaternion groups with amalgamated centre [*10*, p. 238] or the similarly constructed groups of ERDÖS [*4*, p. 58].

We further denote by $[II^*]$ the class of groups

$$G^* = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots\}$$

subject to

$$\begin{cases} [a_i^*, b_i^*] = c^* \neq 1, \\ [a_i^*, a_j^*] = [a_i^*, b_j^*] = [b_i^*, b_j^*] = 1, \\ c^{*p} = 1, \end{cases} \tag{7.3}$$

(where again $i, j = 1, 2, \ldots$ and $i \neq j$) and, moreover,

$$c^* \notin \{a_1^*, a_2^*, \ldots\}.$$

The first three conditions ensure that such a group G^* is a II-group.

Lemma 7.4. *Every group $G \in [II]$ contains a subgroup $G^* \in [II^*]$.*

Proof. We construct G^* by an inductive procedure. Put $c^* = c$, $G_0^* = A_0^* = \{c^*\}$, and assume we have already defined

$$G_n^* = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots, a_n^*, b_n^*\},$$
$$A_n^* = \{a_1^*, a_2^*, \ldots, a_n^*\}$$

such that (7.3) is satisfied for $i, j = 1, \ldots, n$, $i \neq j$, that $c^* \notin A_n^*$, and that

$$A_n^* \subseteq A_{2n} = \{a_1, a_2, \ldots, a_{2n}\}.$$

We apply Lemma 6.4, with $R = A_n^*$, $s = 2n$, $u = 1$, $t = 2$, and $k_1 = c^*$, and find an element

$$a_{n+1}^* = a_{2n+1}^{\lambda(1)} a_{2n+2}^{\lambda(2)}$$

such that

$$c^* \notin A_{n+1}^* = \{a_1^*, a_2^*, \ldots, a_{n+1}^*\}.$$

Not both of $\lambda(1), \lambda(2)$ are divisible by p. Let $\lambda(z)$, where $z = 1$ or 2, be prime to p, and let μ be its reciprocal modulo p:

$$\lambda(z) \cdot \mu \equiv 1 \pmod{p}.$$

Put

$$b_{n+1}^* = b_{2n+z}^{\mu}.$$

Then $[a_{n+1}^*, b_{n+1}^*] = c^*$. It follows that $G_{n+1}^* = \{c^*, a_1^*, b_1^*, \ldots, a_{n+1}^*, b_{n+1}^*\}$ and A_{n+1}^* have again all the required properties. Finally we put

$$G^* = \bigcup_{n=1}^{\infty} G_n^*;$$

then $G^* \in [II^*]$, and the lemma follows.

Lemma 7.5. *If $G^* \in [II^*]$ then $G^* \notin [X]$.*

Proof. We show that all the conjugates $b_i^{*-1} A^* b_i^*$ are distinct. Let $i \neq j$; then $b_i^{*-1} A^* b_i^*$ contains

$$b_i^{*-1} a_i^* b_i^* = a_i^* c^*$$

whereas $b_j^{*-1} A^* b_j^*$ contains

$$b_j^{*-1} a_i^* b_j^* = a_i^*.$$

If $b_i^{*-1} A^* b_i^*$ coincided with $b_j^{*-1} A^* b_j$, then $c^* \in b_i^{*-1} A^* b_i^*$ and also

$$c^* = b_i^* c^* b_i^{*-1} \in b_i^* (b_i^{*-1} A^* b_i^*) b_i^{*-1} = A^*.$$

But this contradicts the definition of a II^*-group. Thus A^* has infinitely many distinct conjugates, and the lemma follows.

Corollary 7.6. $[II] \cap [X] = \emptyset$.

This follows from Lemmas 7.4, 7.5, and 2.2.

§ 8. The class [*III*].

We denote by $[III]$ the class of those I-groups G that satisfy, in addition to (5.1) and (6.1), the conditions

(8.1) $\quad c_i \neq c_j \quad$ when $\quad i \neq j$.

In such a group the derived group

$$C = \{c_1, c_2, \ldots\}$$

is infinite, and using (5.2) we see that

$$[III] \subseteq [FC] - [FD].$$

The existence of such groups is easily verified: the restricted direct product of infinitely many quaternion groups may serve as an example.

Lemma 8.2. *If $G \in [III]$ then G is not a BFC-group.*

Proof. The element $a = a_1 a_2 \ldots a_q$ has at least $q+1$ conjugates, namely apart from itself the q elements

$$b_i^{-1} a\, b_i = a\, c_i \qquad (i = 1, 2, \ldots, q).$$

These are all distinct. As q can be taken arbitrarily large, the cardinals of classes of conjugate elements are unbounded, and the lemma follows.

Next we denote by $[III^*]$ the class of groups $G \in [III]$ that satisfy, moreover,

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{8.3}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$.

In such a group at most one c_i can belong to A, and we may assume that it is c_1, if any: then $c_i \notin A$ for $i = 2, 3, \ldots$.

Lemma 8.4. *A III^*-group G is neither an X-group nor a Y-group.*

Proof. We first show that all the conjugates $b_i^{-1} A\, b_i$ are distinct. Assume on the contrary that $i \neq j$ but $b_i^{-1} A\, b_i = b_j^{-1} A\, b_j$. Then both the elements

$$b_i^{-1} a_i b_i = a_i c_i,$$

$$b_j^{-1} a_i b_j = a_i$$

are contained in $b_i^{-1} A\, b_i$, and so then is c_i. Similarly $c_j \in b_i^{-1} A\, b_i$, and then

$$c_i^{-1} c_j = b_i c_i^{-1} c_j b_i^{-1} \in b_i (b_i^{-1} A\, b_i)\, b_i^{-1} = A,$$

contrary to (8.3). Thus A has infinitely many distinct conjugates, and $G \notin [X]$. Next, a normal subgroup H of G that contains A must also contain all commutators $[a_i, b_i] = c_i$, but by (8.3) no two of these are congruent modulo A, and $|\, H:A \,|$ is infinite. It follows that $G \notin [Y]$. This completes the proof of the lemma.

Lemma 8.5. *Every group $G \in [III]$ contains a subgroup $G^* \in [III^*]$.*

Proof. The plan of the proof is as for Lemma 7.4, but it requires more elaborate arguments. The generators of the group G^* to be constructed will be denoted by $a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots$, and conditions like (5.1), (6.1), when applied to these generators, will also be given an asterisk. Assume the groups

$$G_n^* = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots, a_n^*, b_n^*, c_n^*\},$$

$$A_n^* = \{a_1^*, a_2^*, \ldots, a_n^*\},$$

$$C_n^* = \{c_1^*, c_2^*, \ldots, c_n^*\}$$

are defined and satisfy (5.1*), (6.1*), (8.1*), and (8.3*) as far as applicable; and assume further that there is a number $s = s(n)$ such that

$$A_n^* \leqq A_s = \{a_1, a_2, \ldots, a_s\},$$
$$C_n^* \leqq C_s = \{c_1, c_2, \ldots, c_s\},$$

and if $i \neq j$ and $c_i^{-1} c_j \in A_n^*$ then $i \leq s$ and $j \leq s$. We then determine t such that

$$\frac{p^t - 1}{p - 1} > (n + t)(n + t - 1)$$

and put $(n+t)(n+t-1) = u$. We further put

$$\gamma_1 = c_1^*, \gamma_2 = c_2^*, \ldots, \gamma_n = c_n^*, \gamma_{n+1} = c_{s+1}, \gamma_{n+2} = c_{s+2}, \ldots, \gamma_{n+t} = c_{s+t}$$

and denote the $(n+t)(n+t-1) = u$ quotients $\gamma_i^{-1}\gamma_j$ with $i \neq j$ by $k_1, k_2, \ldots, k_u$. None of these is contained in A_n^*. We can, therefore, apply Lemma 6.4 and Corollary 6.5, with $R = A_n^*$, and find an element

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)}$$

such that $\{A_n^*, a^*\}$ contains none of the elements $k_1, k_2, \ldots, k_u$. Moreover at least one of the exponents $\lambda(1), \ldots, \lambda(t)$, let us say $\lambda(z)$, is not divisible by p. Let μ be its reciprocal modulo the order $p^{\varrho(s+z)}$ of c_{s+z}, that is

$$\lambda(z) \cdot \mu \equiv 1 \pmod{p^{\varrho(s+z)}}.$$

We put

$$a_{n+1}^* = a^*,$$
$$b_{n+1}^* = b_{s+z}^{\mu}$$

and

$$[a_{n+1}^*, b_{n+1}^*] = c_{s+z} = c_{n+1}^*.$$

Then

$$G_{n+1}^* = \{G_n^*, a_{n+1}^*, b_{n+1}^*, c_{n+1}^*\},$$
$$A_{n+1}^* = \{A_n^*, a_{n+1}^*\},$$
$$C_{n+1}^* = \{C_n^*, c_{n+1}^*\}$$

satisfy (5.1*), (6.1*), (8.1*), (8.3*) as far as applicable. Moreover, as A_{n+1}^* is a finitely generated abelian group, it can contain only a finite number of quotients $c_i^{-1} c_j$ for $i \neq j$. There is thus an integer $s' = s(n+1)$ such that if $i \neq j$ and $c_i^{-1} c_j \in A_{n+1}^*$, then $i \leq s'$, $j \leq s'$, and such that at the same time $A_{n+1}^* \leqq A_{s'}$ and $C_{n+1}^* \leqq C_{s'}$. Thus the procedure works inductively; to start it we put $G_0^* = A_0^* = C_0^* = \{1\}$ and $s(0) = 0$. Finally we put

$$G^* = \bigcup_{n=1}^{\infty} G_n^*;$$

then $G^* \in [III^*]$ and the lemma follows.

Corollary 8.6. $[III] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset$.

This follows from Lemmas 8.2, 8.4, 8.5, and 2.2.

§ 9. The class [J].

We denote by $[J]$ the class of H-groups G that satisfy, in addition to (5.1), the relations

$$\begin{cases} [a_i, c_i] = [b_i, c_i] = 1, \\ c_i^{p(i)} = 1, \end{cases} \tag{9.1}$$

where $i = 1, 2, \ldots$ and $p(1), p(2), \ldots$ are different prime numbers.

Again there may be further relations, but they must be compatible with $c_i \neq 1$ for all i. In a J-group G the subgroup

$$C = \{c_1, c_2, \ldots\}$$

is the derived group; it is contained in the centre, in other words G is nilpotent of class 2. If λ, μ are any integers, then

$$[a_i^\lambda, b_i^\mu] = c_i^{\lambda\mu}.$$

Hence a_i (and also b_i) has order $p(i)$ modulo the centre.

Lemma 9.2. *If $G \in [J]$ then G is not a BFC-group.*

Proof. The element $a = a_1 a_2 \ldots a_q$ has at least $q+1$ conjugates, namely apart from itself the q elements

$$b_i^{-1} a\, b_i = a\, c_i \qquad (i = 1, 2, \ldots, q).$$

They are all distinct, because the c_i have different orders and are, therefore, distinct. As q can be taken arbitrarily large, the cardinals of classes of conjugate elements are unbounded, and the lemma follows.

Next we denote by $[J^*]$ the class of groups $G \in [J]$ that satisfy, moreover,

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{9.3}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$.

Lemma 9.4. *A J^*-group G is neither an X-group nor a Y-group.*

The proof is the same as that of Lemma 8.4, with (9.3) substituted for (8.3).

Lemma 9.5. *Let N be a group generated by two elements a, b such that*

$$[a, b] = c \neq 1,$$
$$[a, c] = [b, c] = 1,$$
$$c^p = 1$$

where p is an odd prime. Then N has a subgroup $N^* = \{a^*, b^*\}$ such that the order of a^* is infinite or a power of p, and

$$[a^*, b^*] = c \notin \{a^*\}.$$

Proof. If a has infinite order we put $a^* = a$, $b^* = b$; if a has finite order and b infinite order, we put $a^* = b$, $b^* = a^{-1}$. There only remains the case that a, b have finite orders α, β respectively. The order of N then is finite, namely a divisor of $\alpha\beta p$. We choose numbers λ, μ such that the element

$$a^* = a^\lambda b^\mu$$

(*i*) has order a power of p, (*ii*) does not generate c (that is, $c \notin \{a^*\}$), (*iii*) has maximal order subject to (*i*), (*ii*). Then λ, μ are not both divisible by p; for assume on the contrary that

$$\lambda = p\,\lambda', \quad \mu = p\,\mu'$$

and put

$$a' = a^{\lambda'} b^{\mu'}.$$

Then—as p is an odd prime and therefore divides $\frac{1}{2}p(p-1)$—

$$a'^p = a^{p\lambda'} b^{p\mu'} c^{\frac{1}{2}p(p-1)\lambda'\mu'} = a^\lambda b^\mu = a^*,$$

and a' satisfies (*i*) and (*ii*) and has greater order than a^*, contrary to the choice of a^*. If λ is prime to p, we put $b^* = b^\nu$ where ν is the reciprocal of λ modulo p, that is $\lambda\nu \equiv 1 \pmod{p}$. If λ is divisible by p, we put $b^* = a^\pi$ where similarly $\mu\pi \equiv -1 \pmod{p}$. It is then easy to verify that $N^* = \{a^*, b^*\}$ has the required properties, and the lemma follows.

Lemma 9.6. *Every group $G \in [J]$ has a homomorphic image G^0 with a subgroup $G^* \in [J^*]$.*

Proof. We apply the isoclinic homomorphism of G onto G^0 described in the proof of Lemma 4.7; then G^0 is periodic by Corollary 4.8, and one easily verifies that G^0 is also a J-group. Next, if one of the primes $p(i)$ is 2, we omit the corresponding a_i, b_i, c_i from the generators. We then rename and renumber the generators so that

$$G^0 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\}.$$

In each of the groups $N_i = \{a_i, b_i\}$ we choose a subgroup $N_i^* = \{a_i^*, b_i^*\}$ according to Lemma 9.5; then a_i^*, being periodic, has order a power of $p(i)$, and $[a_i^*, b_i^*] = c_i \notin \{a_i^*\}$. We put

$$G^* = \{a_1^*, b_1^*, c_1, a_1^*, b_1^*, c_2, \ldots\}$$

and

$$A^* = \{a_1^*, a_2^*, \ldots\}, \quad C = \{c_1, c_2, \ldots\}.$$

Then $K^* = CA^*$ is a periodic abelian group. Its primary components are either trivial, namely if the prime p does not occur among the $p(i)$, or the direct product

$$\{c_i\} \times \{a_i^*\},$$

if $p = p(i)$. As a periodic abelian group is the restricted direct product of its primary components, K^* is seen to be the restricted direct product of all cycles $\{c_i\}$ and $\{a_i^*\}$, and then also of C and A^*:

$$K^* = C \times A^*.$$

The defining property (9.3) of a J^*-group is a trivial consequence; thus $G^* \in [J^*]$, and the lemma follows.

Corollary 9.7. $\quad [J] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset.$

This follows from Lemmas 9.2, 9.4, 9.6, and 2.2.

§ 10. The class [K].

We denote by $[K]$ the class of H-groups G in which distinct subgroups $N_i = \{a_i, b_i\}$ have trivial intersection:

$$N_i \cap N_j = \{1\} \quad \text{when} \quad i \neq j. \tag{10.1}$$

It follows immediately from the remarks at the beginning of § 5 that G then is the restricted direct product of $N_1, N_2, \ldots$. Conversely, if an infinite sequence of strictly non-abelian two-generator FC-groups is given, then their restricted direct product is easily seen to be a K-group.

The derived group G' of a K-group G is the restricted direct product of the derived groups $N_1', N_2', \ldots$, none of which is trivial: hence G' is infinite, and using (5.2) we see that

$$[K] \subseteqq [FC] - [FD].$$

It may be remarked that if every N_i, and therefore also G, is nilpotent of class 2, then G is also a III-group; but not every III-group is a K-group.

Lemma 10.2. *If $G \in [K]$ then G is not a BFC-group.*

This is proved as Lemma 8.2.

Next we denote by $[K^*]$ the class of groups $G \in [K]$ that satisfy, moreover,

$$c_i \notin \{a_i\} \tag{10.3}$$

for $i = 1, 2, \ldots$.

Such a group satisfies the apparently stronger condition

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{10.4}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$. For every element of a K-group G is of the form

$$g = n_1 n_2 \ldots n_q, \qquad n_i \in N_i,$$

and this representation is unique apart from trivial factors that can be added or omitted with a change of q. Thus if $c_i^{-1} c_j \in A$, then $c_i^{-1} c_j = a_i^{\lambda} a_j^{\mu}$ and $c_i = a_i^{-\lambda}$, $c_j = a_j^{\mu}$, contrary to (10.3).

Lemma 10.5. *A K^*-group G is neither an X-group nor a Y-group.*

This follows from (10.4) in precisely the same way as Lemma 8.4 follows from (8.3).

Lemma 10.6. *Every group $G \in [K]$ contains a subgroup $G^* \in [K^*]$.*

Proof. If (for some value of i) $c_i \in \{a_i\}$, then b_i transforms a_i into a power of a_i. We therefore have to look for elements that are transformed into other elements not powers of themselves. A non-abelian group in which all the conjugates of every element are amongst its powers only, is hamiltonian; the direct product of two non-abelian groups is not hamiltonian. Hence in every product $N_{2i-1} \times N_{2i}$ we can select two elements a_i^*, b_i^*, say, such that b_i^* transforms a_i^* into an element that is not a power of a_i^*. Then

$$[a_i^*, b_i^*] = c_i^* \notin \{a_i^*\}.$$

Finally we put $G^* = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. This is then a K^*-group, and the lemma follows.

Corollary 10.7. $[K] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset$.

This follows from Lemmas 10.2, 10.5, 10.6, and 2.2.

§ 11. The class $[FC] - [FD]$.

Throughout this section we assume that G is an FC-group but not an FD-group. By selecting suitable subgroups we show that G is then neither a BFC-group, nor an X-group, nor a Y-group.

Lemma 11.1. *Every group $G \in [FC] - [FD]$ contains a subgroup $G^1 \in [H] - [FD]$.*

Proof. We construct G^1 by an inductive process. Assume elements $a_1, b_1, c_1, \ldots, a_{n-1}, b_{n-1}, c_{n-1}$ so chosen in G that

$$\text{(11.2)} \qquad \begin{cases} [a_i, b_i] = c_i \neq c_j \neq 1, \\ [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1 \end{cases}$$

for all i and $j \neq i$. This is vacuously true for $n = 1$. The centralizer $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ has finite index in G, and cannot be an FD-group, by Corollary 4.2. The derived group of $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ is then infinite, and we can find two elements a_n, b_n, say, in $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ whose commutator $c_n = [a_n, b_n]$ is different from $1, c_1, c_2, \ldots, c_{n-1}$. Then $a_1, b_1, c_1, \ldots, a_n, b_n, c_n$ clearly satisfy (11.2). Finally we put

$$G^1 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\};$$

then G^1 is seen to be an H-group, and the construction ensures that the commutators c_i are all different; hence $G^1 \notin [FD]$, and the lemma follows.

Lemma 11.3. *If either $[a_i, c_i] \neq 1$ or $[b_i, c_i] \neq 1$ for infinitely many values of i, then G^1 is neither a BFC-group, nor an X-group, nor a Y-group.*

Proof. Let G^2 denote the subgroup of G^1 that is generated by all $N_i = \{a_i, b_i\}$ $(= \{a_i, b_i, c_i\})$ in which a_i, b_i do not both commute with c_i. Then G^2 is again an H-group but not an FD-group. Denote the centre of G^2 by Z^2. Its factor group G^2/Z^2 is the restricted direct product of infinitely many finite non-abelian groups

$$F_i = N_i Z^2/Z^2,$$

and is, therefore, a K-group. Then by Corollary 10.7 and repeated application of Lemma 2.2, $G^1 \notin [BFC] \cup [X] \cup [Y]$, and the lemma follows.

Lemma 11.4. *If $[a_i, c_i] = [b_i, c_i] = 1$ for all but a finite number of values of i, then G^1 contains a subgroup $G^3 \in [H] - [FD]$ that is nilpotent of class 2 (or, in other words, all of whose commutators lie in the centre).*

This is trivial: we only have to omit from the set of generators those finitely many a_i, b_i, c_i for which $[a_i, c_i] \neq 1$ or $[b_i, c_i] \neq 1$.

We note that in such a nilpotent group of class 2

$$[a, b] = c \quad \text{implies} \quad [a^\lambda, b^\mu] = c^{\lambda\mu} \tag{11.5}$$

—a fact that we have repeatedly used. Next we note that the derived group

$$C^3 = \{c_1, c_2, \ldots\}$$

of G^3 is, as a periodic abelian group, the restricted direct product of its primary components.

Lemma 11.6. *Let $G^3 \in [H] - [FD]$ be nilpotent of class 2. If its derived group C^3 has an infinite primary component, then G^3 contains a subgroup $G^4 \in [III]$.*

Proof. Let the primary p-component C_p^3 of C^3 be infinite. If the order $\gamma(i)$ of c_i is factorized in the form

$$\gamma(i) = p^{\varrho(i)}\gamma'(i)$$

with $\gamma'(i)$ prime to p, then

$$C_p^3 = \{c_1^{\gamma'(1)}, c_2^{\gamma'(2)}, \ldots\}.$$

Also, by (11.5)

$$c_i^{\gamma'(i)} = [a_i, b_i^{\gamma'(i)}].$$

Among the $c_i^{\gamma'(i)}$ there must be infinitely many that are different from one another; we choose accordingly an infinite sequence $i(1), i(2), \ldots$ such that when $j \neq k$ then

$$c_j^* = c_{i(j)}^{\gamma'(i(j))} \neq c_k^* = c_{i(k)}^{\gamma'(i(k))},$$

and put

$$a_j^* = a_{i(j)}, \qquad b_j^* = b_{i(j)}^{\gamma'(i(j))}.$$

Finally we put $G^4 = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. It is then immediately verified that G^4 is a III-group, and the lemma follows.

Lemma 11.7. *Let $G^3 \in [H] - [FD]$ be nilpotent of class 2. If its derived group C^3 has no infinite primary component, then G^3 contains a subgroup $G^5 \in [J]$.*

Proof. As C^3 is infinite, it must have infinitely many non-trivial primary components. It follows that there are infinitely many prime numbers each of which divides the order $\gamma(i)$ of at least one c_i; and infinitely many different c_i are required for this. Let then $i(1) < i(2) < \cdots$ be a sequence of indices and $p(1), p(2), \ldots$ a sequence of distinct primes such that $p(j)$ divides $\gamma(i(j))$. We put $\gamma(i(j)) = p(j) \cdot \gamma'(i(j))$ and

$$c_j^* = c_{i(j)}^{\gamma'(i(j))}.$$

Then c_j^* has order $p(j)$. Next we note that by (11.5)

$$c_j^* = [a_{i(j)}, b_{i(j)}^{\gamma'(i(j))}].$$

We therefore put

$$a_j^* = a_{i(j)}, \qquad b_j^* = b_{i(j)}^{\gamma'(i(j))},$$

and finally $G^5 = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. One confirms immediately that G^5 is a J-group, and the lemma follows.

Corollary 11.8. *If $G^3 \in [H] - [FD]$ is nilpotent of class 2, then it is neither a BFC-group, nor an X-group, nor a Y-group.*

This follows from Lemmas 11.6, 11.7, 2.2 and Corollaries 8.6, 9.7.

Corollary 11.9. *If G is an FC-group but not an FD-group, then G is neither a BFC-group, nor an X-group, nor a Y-group.*

This follows from Lemmas 11.1, 11.3, 11.4, 2.2, and Corollary 11.8.

§ 12. The class [*FD*] — [*FIZ*].

We now assume that G is an FD-group but not a FIZ-group and show that it is then not an X-group either.

Lemma 12.1. *Every group $G \in [FD] - [FIZ]$ contains a subgroup $G^1 \in [H]$.*

Proof. We construct G^1 by an inductive process. Assume elements $a_1, b_1, c_1, \ldots, a_{n-1}, b_{n-1}, c_{n-1}$ so chosen that

$$\begin{cases} [a_i, b_i] = c_i \neq 1, \\ [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1 \end{cases} \tag{12.2}$$

for all i and $j \neq i$. This is vacuously true for $n = 1$. The centralizer $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ has finite index in G, and cannot be an FIZ-group, by Corollary 4.4. Hence $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ is *a fortiori* non-abelian, and we can find two elements a_n, b_n, say, in $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ whose commutator $c_n = [a_n, b_n]$ is different from 1. Then $a_1, b_1, c_1, \ldots, a_n, b_n, c_n$ clearly satisfy (12.2). Finally we put

$$G^1 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\};$$

this is clearly an H-group, and the lemma follows. It may be noted that G^1, as a subgroup of an FD-group, is itself an FD-group.

Lemma 12.3. *Every group $G^1 \in [FD] \cap [H]$ contains a subgroup $G^2 \in [II]$.*

Proof. As the derived group of G^1 is finite, only a finite number of the c_i can be distinct, and at least one of them, c say, occurs infinitely often. Let $c = c_{i(1)} = c_{i(2)} = \cdots$. Then c lies in the centre of G^1, being generated by $a_{i(1)}$, $b_{i(1)}$ as well as by $a_{i(2)}$, $b_{i(2)}$ (cf. a remark on the immediate consequences of the definition of an H-group in § 5). As before we then have

$$[a_{i(j)}^{\lambda}, b_{i(j)}^{\mu}] = c^{\lambda\mu}.$$

Let p be a prime number dividing the order γ of c, and let $\gamma = p\,\gamma'$. Then $c^* = c^{\gamma'}$ has order p. We put

$$a_j^* = a_{i(j)}, \quad b_j^* = b_{i(j)}^{\gamma'},$$

and note that

$$[a_j^*, b_j^*] = c^* \neq 1,$$
$$[a_j^*, c^*] = [b_j^*, c^*] = 1,$$
$$c^{*p} = 1.$$

Hence the group $G^2 = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots\}$ is a II-group, and the lemma follows.

Corollary 12.4. *If G is an FD-group but not an FIZ-group, then G is not an X-group either.*

This follows from Lemmas 12.1, 12.3, 2.2 and Corollary 7.6.

§ 13. Conclusion.

We are now in a position to prove the main results announced in the introduction.

Theorem 13.1. *The classes $[BFC]$, $[FD]$, $[Y]$, and $[BY]$ all coincide.*

Proof. By Lemmas 2.3, 3.7

$$[FD] \leqq [BY] \leqq [Y] \leqq [FC].$$

Moreover $[FD] \leqq [BFC] \leqq [FC]$, as the order of the derived group is easily seen to provide a bound for the number of distinct conjugates of an element[5]). On the other hand, by Corollary 11.9,

$$[FC] - [FD] \leqq [FC] - ([BFC] \cup [Y]),$$

whence

$$[FD] = [BY] = [Y] = [BFC],$$

as asserted.

Theorem 13.2. *The classes $[FIZ]$, $[X]$, and $[BX]$ coincide.*

Proof. By Lemmas 2.3, 3.2

$$[FIZ] \leqq [BX] \leqq [X] \leqq [FC].$$

On the other hand, by Corollary 11.9

$$[FC] - [FD] \leqq [FC] - [X],$$

and by Corollary 12.4 also

$$[FD] - [FIZ] \leqq [FD] - [X].$$

Hence

$$[FIZ] = [BX] = [X],$$

as asserted.

It may be remarked that our methods do not allow us to state a bound for the order of the derived group in a group in which a finite bound is given for the cardinals of the classes of conjugate elements, or for the indices of subgroups in their normal closures; nor can we derive a bound for the index of the centre in a group in which a finite bound is given for the cardinals of the classes of conjugate subgroups. One is naturally led to conjecture that such bounds for the order of the derived group or the index of the centre do exist; and in the case of a given bound for the cardinals of the classes of conjugates elements this has very recently been confirmed by JAMES WIEGOLD [*11*].

---

[5]) Cf. [*9*], proof of Theorem 5.4.

References.

[1] BAER, R.: Finiteness properties of groups. Duke Math. J. **15**, 1021—1032 (1948). — [2] BAER, R.: Das Hyperzentrum einer Gruppe. III. Math. Z. **59**, 299—338 (1953). — [3] DEDEKIND, R.: Über Gruppen, deren sämmtliche Teiler Normalteiler sind. Math. Ann. **48**, 548—561 (1897). — [4] ERDÖS, J.: The theory of groups with finite classes of conjugate elements. Acta Math. Acad. Sci. Hungar. **5**, 45—58 (1954). — [5] HAIMO, F.: Groups with a certain condition on conjugates. Canad. J. Math. **4**, 369—372 (1952). — [6] HAIMO, F.: The *FC*-chain of a group. Canad. J. Math. **5**, 498—511 (1953). — [7] HALL jr., M.: Subgroups of finite index in free groups. Canad. J. Math. **1**, 187—190 (1949). — [8] HALL, P.: The classification of prime-power groups. J. reine angew. Math. **182**, 130—141 (1940). — [9] NEUMANN, B. H.: Groups with finite classes of conjugate elements. Proc. London Math. Soc. (3) **1**, 178—187 (1951). — [10] NEUMANN, B. H.: Groups covered by permutable subsets. J. London Math. Soc. **29**, 236—248 (1954). — [11] WIEGOLD, J.: Unpublished.

Dept. of Math., The University, Manchester 13 (England).

(Eingegangen am 3. Januar 1955.)

ROGOSINSKI, W. W.
Math. Zeitschr. Bd. 63, S. 97—108 (1955)

On finite systems of linear equations with an infinity of unknowns.

Dedicated to the memory of I. SCHUR, the great master and a helpful friend.

By
W. W. ROGOSINSKI.

Introduction.

0.1. We are given n linear equations

$$\sum_{k=0}^{\infty} a_{ik} x_k = C_i \qquad (1 \leq i \leq n) \tag{0.1.1}$$

in an infinity of unknowns x_k. These and the given a_{ik} and C_i are complex (or real) numbers. The n sequences $\{a_{ik}\}$ are supposed to be linearly independent. It is further assumed that either, for a given p with $p \geq 1$,

$$\sum_{k=0}^{\infty} |a_{ik}|^p < \infty \qquad (1 \leq i \leq n); \tag{0.1.2}$$

or [this is the "convergency case" (c)] that

$$\lim_{k \to \infty} a_{ik} = a_{i0} \qquad (1 \leq i \leq n). \tag{0.1.3}$$

Note that, in the case (c), the first term a_{i0} of the sequence $\{a_{ik}\}$ is the limit of this sequence as $k \to \infty$.

0.2. As for the admissible solutions $\{x_k\}$, we demand, when $p > 1$, that

$$\sum_{k=0}^{\infty} |x_k|^{p'} < \infty \quad \text{where} \quad 1/p + 1/p' = 1; \tag{0.2.1}$$

when $p = 1$, that

$$\sup |x_k| < \infty; \tag{0.2.2}$$

and, in the case (c), that

$$\sum_{k=0}^{\infty} |x_k| < \infty. \tag{0.2.3}$$

In all cases, the left hand sides of (0.1.1) will then converge absolutely.

Amongst the infinity of solutions $\{x_k\}$ we consider *minimal solutions* $\{x_k^*\}$. These are solutions for which the left hand sides of (0.2.1), (0.2.2), or (0.2.3), respectively, become minimal. The existence of minimal solutions is in each case easy to prove. Also, if $p > 1$, the nature of the minimal solutions is known[1]). It can be expressed by the following theorem.

[1]) RIESZ, F. [3], Chapter III, p. 42—77; in particular, p. 52—53, formula (21). The numbers refer to the list of references.

Theorem A. *If $p>1$, then the system* (0.1.1) *has exactly one minimal solution* $\{x_k^*\}$. *Moreover,*

$$x_k^* = \operatorname{sign} \overline{Q}_k \, |Q_k|^{p-1} \qquad (k \geq 0)^2), \tag{0.2.4}$$

where the Q_k are of the form

$$Q_k = \sum_{i=1}^{n} A_i a_{ik} \qquad (k \geq 0). \tag{0.2.5}$$

If $p=1$, and in the case (c), an infinity of minimal solutions will usually exist. The following results concerning their nature seem to be new.

Theorem B. *If $p=1$, then there exists at least one index K such that x_K^* is the same for every minimal solution* $\{x_k^*\}$ *of* (0.1.1). *Moreover,*

$$|x_K^*| = \max_k |x_k^*| \tag{0.2.6}$$

for every minimal solution.

In order to state the result in the case (c), we introduce the following terminology. A sequence $\{x_k\}$ is said to be *directed* if either (i) the x_k vanish eventually, or (ii) $\arg x_k \to \varphi \pmod{2\pi}$ as $k \to \infty$ for those k for which $x_k \neq 0$. Two sequences $\{x_k\}$ and $\{\xi_k\}$ are said to be *equally directed* if (α) each is directed, (β) $\arg x_k = \arg \xi_k$ whenever both $x_k \neq 0$ and $\xi_k \neq 0$, and (γ) both sequences have the same "direction" φ when they are in the above case (ii).

Theorem C. *In the case* (c) *every minimal solution* $\{x_k^*\}$ *of* (0.1.1) *is directed, and all are equally directed.*

In the case (c_0), when

$$\lim_{k \to \infty} {}_k = 0 \qquad (1 \leq i \leq n), \tag{0.2.7}$$

so that $a_{i0}=0$, we can amplify Theorem C a little.

Theorem C_0. *In the case* (c_0) *all minimal solutions* $\{x_k^*\}$ *of* (0.1.1) *are equally directed. Moreover, there exists a K, the same for all minimal solutions, such that $x_k^*=0$ when $k>K$.*

The 'real' case of (c), when all a_{ik}, C_i, and x_k are supposed to be real, deserves special mentioning. We shall say that two real sequences $\{x_k\}$ and $\{\xi_k\}$ have *equal sign* if, for every k, x_k and ξ_k are simultaneously non-negative, or non-positive. Theorem C then takes the following form.

Theorem C_r. *In the real case* (c) *all minimal solutions* $\{x_k^*\}$ *of* (0.1.1) *have equal sign. Moreover, there exists a K such that this sign remains constant when $k>K$.*

0.3. The preceding theorems may be illustrated by a simple geometrical situation in Euclidean space which they generalize. Consider, for instance, one real linear equation

$$a x + b y = C \tag{0.3.1}$$

[2]) We write $\operatorname{sign} 0 = 0$ and $\operatorname{sign} z = e^{i\varphi}$, when $\arg z = \varphi$.

in two real unknowns x and y. It may be interpreted as the equation of a line L in the Cartesian x, y-plane. We may also write (0.3.1) in 'infinite' form

$$a\,x + b\,y + o\cdot z + o\cdot s + \cdots = C. \tag{0.3.2}$$

If $p > 1$, it is clear, in view of (0.2.1), that we must have $z^* = s^* = \cdots = 0$ while $|x^*|^{p'} + |y^*|^{p'}$ must be made a minimum for every minimal solution. The simplest case $p = 2$, when $p' = 2$, is typical. The unique minimal solution is given by the point (x^*, y^*) on L nearest to the origin [compare Theorem A].

If $p = 1$, we must make $\max(|x^*|, |y^*|)$ a minimum μ, say. The $z^*, s^*, \ldots$ can be chosen arbitrarily but with absolute values at most μ. Geometrically, we have to find the square $|x| = \mu$, $|y| = \mu$ that touches the line L. There are two cases. Either L passes through exactly one vertex; then this vertex gives the unique minimal values x^* and y^*. Or L has one side in common with the square, when every point of this side gives a minimal pair (x^*, y^*). In either case, one of the coordinates at least is fixed for every minimal solution, with absolute value μ [compare Theorem B].

In the case (c_0), we must again have $z^* = s^* = \cdots = 0$ while $|x^*| + |y^*|$ becomes a minimum μ. This time we have to find the square $|x| + |y| = \mu$ that touches L. We have the same two cases as for $p = 1$. In the first, (x^*, y^*) is unique; in the second, all minimal pairs (x^*, y^*) belong to the same quadrant. In either case, Theorem C_0 is, clearly, satisfied.

0.4. The study of an infinite system of linearly independent linear equations

$$\sum_{k=0}^{\infty} a_{ik}\, x_k = C_i \qquad (i = 1, 2, \ldots) \tag{0.4.1}$$

can be attached, by well known arguments, to that of the finite case[3]). In the first place, of course, the system should have solutions. If so, then, for $p > 1$, there exists exactly one minimal solution, and an analogue of Theorem A remains valid. On the other hand, we shall see that neither Theorem B nor Theorem C holds in the infinite case except under further drastic assumptions.

0.5. The treatment of a finite system of linear equations in an infinity of unknowns in the older literature, as set out, for instance, in the book of F. RIESZ [3], is direct and uses elementary algebraic and variational arguments. Prof. Rado has now also obtained a similar direct proof for our Theorem B. As for Theorem C, the property (β) above of all minimal solutions is quite easy to establish, and it should be possible to obtain also the properties (α) and (γ) by a direct method. My own proofs for all the above theorems are based on the methods of functional analysis applied to the appropriate sequential spaces to which the given $\{a_{ik}\}$ belong. In fact, these theorems turn out to be rather special applications of more general results concerning continuous linear functionals on these spaces, results which deserve some interest of their own. We require the well known form of these functionals and further the familiar HAHN-BANACH extension theorem for functionals. It should be noted that the proof of this theorem requires no transfinite induction since

[3]) RIESZ, F [3], Chapter III.

all our spaces are separable. Our method of proof has the great advantage over the direct and apparently more elementary methods that it covers all cases uniformly in a short and straightforward way, and it seems to me the natural approach to these and many similar extremum problems. In any case, it is hoped that the results of this paper will add another example of the usefulness and power of abstract methods in quite concrete problems of analysis.

1. The general argument.

1.1. Let $\mathfrak{A}$ be a normed linear vector space over the field of complex (or real) numbers; $\mathfrak{A}$ need not be complete (a Banach space). If $a \in \mathfrak{A}$, the norm of a is denoted by $\|a\|$. The continuous (bounded) linear functionals (c.l.f.) on $\mathfrak{A}$ are denoted by $B(a)$; in the complex case they are complex-valued, in the real case real-valued. They form the conjugate space $\mathfrak{B}$ with norm $\|B\|$.

Suppose now that we are given n linearly independent points $a_i \in \mathfrak{A}$ and n arbitrary complex (real) numbers C_i. We may ask for all those B that satisfy

$$B(a_i) = C_i \qquad (1 \leq i \leq n). \tag{1.1.1}$$

Consider the n-dimensional subspace $\mathfrak{P}_n$ of $\mathfrak{A}$ of all linear combinations $p = p_n = \sum_1^n \lambda_i a_i$, where the λ_i are complex (real) numbers. Any solution of (1.1.1) must satisfy

$$B(p) = \sum_1^n \lambda_i B(a_i) = \sum_1^n \lambda_i C_i. \tag{1.1.2}$$

Conversely, the right hand side of (1.1.2) is a c.l.f. $\tilde{B}(p)$ defined on $\mathfrak{P}_n$ [4]) which satisfies the equations

$$\tilde{B}(a_i) = C_i \qquad (1 \leq i \leq n). \tag{1.1.3}$$

The solutions B of (1.1.1) are, therefore, the 'extensions' of $\tilde{B}$ from $\mathfrak{P}_n$ to the whole of $\mathfrak{A}$, i.e. they are those B on $\mathfrak{A}$ for which $B(p) = \tilde{B}(p)$ for all $p \in \mathfrak{P}_n$. If $\|\tilde{B}\|_*$ denotes the norm of $\tilde{B}$ on $\mathfrak{P}_n$, then, clearly,

$$\|B\| \geq \|\tilde{B}\|_* \tag{1.1.4}$$

for any such extension. We now apply the HAHN-BANACH extension theorem [5]) to the functionals $\tilde{B}$ on $\mathfrak{P}_n$. According to this theorem there exists at least one extension B^* of $\tilde{B}$ from $\mathfrak{P}_n$ to $\mathfrak{A}$ that is minimal in the sense that

$$\|B^*\| = \|\tilde{B}\|_*. \tag{1.1.5}$$

In view of (1.1.4) we call such a B^* a *minimal solution* for the system (1.1.1).

[4]) See [6], p. 213.

[5]) See [1], p. 55; and [2].

Next, it follows from the form (1.1.2) of $\widetilde{B}$, that there exists at least one point $P=\sum_1^n \Lambda_i a_i$ of $\mathfrak{P}_n$ such that $\|P\|=1$ and

$$\widetilde{B}(P)=\sum_1^n \Lambda_i C_i=\|\widetilde{B}\|_*. \tag{1.1.6}$$

We call P a *maximal point* for $\widetilde{B}$. By (1.1.5), we have

$$B^*(P)=\|B^*\|. \tag{1.1.7}$$

We note that this relation must hold for *every* pair of a minimal B^* and a maximal P. For any $p\in\mathfrak{P}_n$ with $\|p\|=1$ we have, of course, $|B^*(p)|\leq\|B^*\|$. *The discussion of equality here, as it holds in* (1.1.7), *allows in many cases additional information about the functional* B^*. This simple remark provides a powerful argument in applications.

We add the following remarks. The set of all minimal extensions B^* of $\widetilde{B}$ is a convex set in $\mathfrak{B}$. Similarly, the set of all maximal points P is convex in $\mathfrak{P}_n$, provided, however, that $\|\widetilde{B}\|_*>0$. In the "trivial" case when $\|\widetilde{B}\|_*=0$, all $p\in\mathfrak{P}_n$ with $\|p\|=1$ are maximal[6]).

A normed linear vector space $\mathfrak{B}$ is said to be *strictly convex* if, in the triangle relation

$$\|v_1+v_2\|\leq\|v_1\|+\|v_2\| \tag{1.1.8}$$

for the norm, equality occurs if, and only if, $\lambda_1 v_1=\lambda_2 v_2$ where $\lambda_1\geq 0$, $\lambda_2\geq 0$ and $\lambda_1+\lambda_2>0$.

If $\mathfrak{B}$ is strictly convex then the above convex set of the B^* reduces to one point: the minimal solution of (1.1.1) is then unique. Similarly, if $\mathfrak{P}_n$ is strictly convex, the maximal point P for $\widetilde{B}$ is unique, provided that $\|\widetilde{B}\|_*>0$ [6]).

1.2. We consider, next, an infinite system of equations

$$B(a_i)=C_i \qquad (i\geq 1). \tag{1.2.1}$$

We assume that the sequence of points $a_i\in\mathfrak{A}$ is linearly independent. Let B_n^* be a minimal solution for the corresponding finite system (1.1.1). Clearly, the $\|B_n^*\|$ increase with n. It is well known that the system (1.2.1) has solutions $B\in\mathfrak{B}$ if, and only if,

$$\|B_n^*\|\uparrow M, \tag{1.2.2}$$

where M is finite[7]). Moreover, if this condition is satisfied, then there exists at least one minimal solution B^* of (1.2.1) for which $\|B^*\|=M$. For all other solutions, $\|B\|\geq M$. Again, the minimal solutions form a convex set in $\mathfrak{B}$, and if $\mathfrak{B}$ is strictly convex, then B^* is unique.

---

[6]) [5], p. 299—300.

[7]). [1], Theorem 5, p. 57 (for a real space).

Let $\mathfrak{P}=\bigcup_1^\infty \mathfrak{P}_n$ and let $\overline{\mathfrak{P}}$ be the closure of $\mathfrak{P}$ in $\mathfrak{A}$. If the unit sphere of $\overline{\mathfrak{P}}$ is weakly compact (or, equivalently, if $\overline{\mathfrak{P}}$ is reflexive), then there exists at least one maximal point $P\in\overline{\mathfrak{P}}$, such that $\|P\|=1$ and

$$B^*(P)=\|B^*\|. \tag{1.2.3}$$

In this case we may again be able to draw some additional information concerning the minimal solutions B^*. It should, however, be noted, that without such or similar assumptions there will be, in general, no maximal points P to give us this information. Lastly, we remark that there exists at most one maximal point if $\overline{\mathfrak{P}}$ is strictly convex, provided that $\|B^*\|>0$.

2. The space l^p, $p>1$.

2.1. We consider the Banach space l^p, $p>1$, of all complex (real) sequences $a=\{a_k\}$, $k\geq 0$, with $\sum_{k=0}^{\infty}|a_k|^p<\infty$. The norm is $\|a\|_p=\left[\sum_{k=0}^{\infty}|a_k|^p\right]^{1/p}$.

It is well known that every c.l.f. b on l^p is of the form[8])

$$b(a)=\sum_{k=0}^{\infty}a_k x_k, \tag{2.1.1}$$

where

$$x=\{x_k\}\in l^{p'}, \qquad 1/p+1/p'=1, \tag{2.1.2}$$

and $\|b\|=\|x\|_{p'}$. We call x the *kernel* of b; it is uniquely determined by b.

2.2. Let $\mathfrak{g}$ be a linear subspace of l^p. If $\tilde{b}$ is a c.l.f. defined on $\mathfrak{g}$, with norm $\|\tilde{b}\|_*$ on $\mathfrak{g}$, then, by the HAHN-BANACH extension theorem, there exists a representation

$$\tilde{b}(g)=\sum_{k=0}^{\infty}g_k x_k^* \qquad (g=\{g_k\}\in\mathfrak{g}) \tag{2.2.1}$$

where

$$x^*=\{x_k^*\}\in l^{p'} \quad \text{and} \quad \|x^*\|_{p'}=\|\tilde{b}\|_*. \tag{2.2.2}$$

In general, $\tilde{b}$ will also have other representations, with other kernels $x\in l^{p'}$. Then $\|x\|_{p'}\geq\|\tilde{b}\|_*$ for all such kernels, so that x^* appears as a *minimal kernel* for $\tilde{b}$. Since the conjugate space $l^{p'}$ of l^p (the space of the c.l.f. on l^p) is strictly convex (by Minkowski's inequality), *the minimal kernel x^* is unique.*

2.3. A sequence $G=\{G_k\}\in\mathfrak{g}$ is said to be *maximal* for $\tilde{b}$ if

$$\text{(i)}\ \ \|G\|_p=1, \qquad \text{(ii)}\ \ \tilde{b}(G)=\|\tilde{b}\|_*. \tag{2.3.1}$$

If $\|\tilde{b}\|_*=0$, then all $G\in\mathfrak{g}$ with $\|G\|_p=1$ are maximal. If $\|\tilde{b}\|_*>0$ and $\mathfrak{g}$ is closed, then there exists exactly one maximal G. For, l^p is both reflexive

[8]) [1], p. 68 (for the real space l^p).

(its unit sphere is weakly compact) and strictly convex. In this case, we conclude from

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x_k^* = \|\tilde{b}\|_* = \|x^*\|_{p'} \tag{2.3.2}$$

that

$$x_k^* = \operatorname{sign} \overline{G}_k \, |G_k|^{p-1} \|\tilde{b}\|_* . \tag{2.3.3}$$

We thus obtain the following result.

Theorem 1. *Let $p > 1$, and let $\mathfrak{g}$ be a closed subspace of $\mathfrak{l}^p$. If $\tilde{b}$ is a continuous linear functional on $\mathfrak{g}$, then there exists a unique $\Gamma = \{\Gamma_k\} \in \mathfrak{g}$ such that*

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k \operatorname{sign} \overline{\Gamma}_k \, |\Gamma_k|^{p-1}; \quad \|\tilde{b}\|_* = \|\Gamma\|_p^{p-1}. \tag{2.3.4}$$

Remarks. (i) If $\|\tilde{b}\|_* = 0$, then $\Gamma = 0$; otherwise $\Gamma = \|\tilde{b}\|_*^{1/p-1} G$. (ii) If $p = 2$, then there exists a unique $\Gamma \in \mathfrak{g}$ such that

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k \overline{\Gamma}_k; \quad \|\tilde{b}\|_* = \|\Gamma\|_2. \tag{2.3.5}$$

(iii) If $\mathfrak{g} = \mathfrak{l}^p$, we find the existence of exactly one $A = A_k \in \mathfrak{l}^p$ such that

$$b(a) = \sum_{k=0}^{\infty} a_k \operatorname{sign} \overline{A}_k \, \|A_k\|^{p-1}; \quad \|b\| = \|A\|_p^{p-1}. \tag{2.3.6}$$

2.4. We use the notations in § 1. Let the $a_i = \{a_{ik}\}$, $1 \leq i \leq n$, $k \geq 0$, belong to $\mathfrak{A} = \mathfrak{l}^p$. By (2.1.1), the system of equations (1.1.1) in the unknown functionals $B = b$ takes the form of the system of linear equations (0.1.1) in the unknowns x_k. In fact, the conditions (0.1.2) and (0.2.1) state that $a_i \in \mathfrak{l}^p$ and $x = \{x_k\} \in \mathfrak{l}^{p'}$. We now apply Theorem 1 to the closed subspace $\mathfrak{g} = \mathfrak{P}_n$ spanned by the a_i and obtain Theorem A of the introduction.

We note also that Theorem A remains valid, with

$$x_k^* = \operatorname{sign} \overline{\Gamma}_k \, |\Gamma_k|^{p-1}, \quad \Gamma \in \overline{\mathfrak{P}}, \tag{2.4.1}$$

for an infinite system (0.4.1) of linear equations, provided that this system has at all solutions $x \in \mathfrak{l}^{p'}$. The necessary and sufficient condition for this is that the sequence of the $\|x^*(n)\|_p$ should remain bounded, where $x^*(n)$ is the minimal solution of the corresponding finite system (0.1.1).

3. The space $\mathfrak{l}$.

3.1. We consider the Banach space $\mathfrak{l}$ of all complex (real) sequenses $a = \{a_k\}$, $k \geq 0$, with $\sum_{k=0}^{\infty} |a_k| < \infty$. The norm is $\|a\| = \sum_{k=0}^{\infty} \|a_k\|$. Every c.l.f. b on $\mathfrak{l}$ can be represented in the form[9])

$$b(a) = \sum_{k=0}^{\infty} a_k x_k \tag{3.1.1}$$

[9]) [1], p. 67 (for the real space $\mathfrak{l}$).

where

$$\|b\| = \sup |x_k| \quad (<\infty). \tag{3.1.2}$$

The *kernel* $x = \{x_k\}$ is uniquely determined by b.

3.2. Let $\mathfrak{g}$ be a linear subspace of $\mathfrak{l}$. If $\tilde{b}$ is a c.l.f. defined on $\mathfrak{g}$, with norm $\|\tilde{b}\|_*$ on $\mathfrak{g}$, then, by the Hahn-Banach theorem, there exists a representation

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k x_k^* \qquad (g = \{g_k\} \in \mathfrak{g}) \tag{3.2.1}$$

where

$$\|\tilde{b}\|_* = \sup |x_k^*|. \tag{3.2.2}$$

Again $x^* = \{x_k^*\}$ appears as a *minimal kernel* for $\tilde{b}$ in the sense that, for every other representation of $\tilde{b}$ with kernel $x = \{x_k\}$, $\sup |x_k| \geq \|\tilde{b}\|_*$. The conjugate space of $\mathfrak{l}$ is the space $\mathfrak{l}^\infty$ of all bounded sequences $x = \{x_k\}$, with norm $\|x\| = \sup |x_k|$. It is not strictly convex. Hence, in general, we shall have an infinity of minimal kernels x^* for $\tilde{b}$. They form a convex set in $\mathfrak{l}^\infty$, lying on the sphere $\|x\| = \|\tilde{b}\|_*$.

3.3. A sequence $G = \{G_k\} \in \mathfrak{g}$ is said to be *maximal* for $\tilde{b}$ if

$$\text{(i)} \quad \|G\| = 1, \qquad \text{(ii)} \quad \tilde{b}(G) = \|\tilde{b}\|_*. \tag{3.3.1}$$

If $\|\tilde{b}\|_* = 0$, then all $G \in \mathfrak{g}$ with $\|G\| = 1$ are maximal. If $\|\tilde{b}\|_* > 0$, there may be no maximal sequences, even when $\mathfrak{g}$ is closed [since $\mathfrak{l}$ is not reflexive]; and there may be many [since $\mathfrak{l}$ is not strictly convex]. They form a convex set lying on the unit sphere $\|g\| = 1$ of $\mathfrak{g}$.

We now *assume* that $\tilde{b}$ has a maximal G. Then

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x_k^* = \|\tilde{b}\|_* = \sup |x_k^*|. \tag{3.3.2}$$

It, clearly, follows that

$$x_k^* = \operatorname{sign} \overline{G}_k \|\tilde{b}\|_* \quad \text{whenever} \quad G_k \neq 0. \tag{3.3.3}$$

We note that the set $K(G) = [k : G_k \neq 0]$ is not empty since $\|G\| = 1$. Also (3.3.3) must hold for every pair of a maximal G and a minimal x^*. It follows that, for all minimal kernels x^*, the x_k^* with k belonging to the non-empty "index set" $K = \bigcup_G K(G)$ are the same, and that for them $|x_k^*| = \|\tilde{b}\|_*$.

Theorem 2. *Suppose that $\mathfrak{g}$ is a subspace of $\mathfrak{l}$ and that the continuous linear functional $\tilde{b}$ on $\mathfrak{g}$ has a maximal sequence G. Then there exists a non-empty index set K and a uniquely determined sequence $\{\eta_k\}$ for $k \in K$, with $|\eta_k| = \|\tilde{b}\|_*$, such that every minimal representation of $\tilde{b}$ is of the form*

$$\tilde{b}(g) = \sum_{k \in K} g_k \eta_k + \sum_{k \notin K} g_k x_k^* \tag{3.3.4}$$

where $|x_k^*| \leq \|\tilde{b}\|_*$.

3.4. It follows from (3.3.4) that a necessary condition for $\tilde{b}$ to have a maximal G is that *all* minimal kernels x^* should satisfy

$$\sup|x_k^*| = \max|x_k^*| = \|\tilde{b}\|_*. \tag{3.4.1}$$

Conversely, suppose that *some* minimal kernel x^{**} satisfies (3.4.1). Let $K^* = [k:|x_k^{**}| = \|\tilde{b}\|_*]$. Then any sequence G with $\|G\| = 1$, satisfying

$$G_k = \begin{cases} 0 \\ \text{or} \quad x_k^{**} = \|\tilde{b}\|_* \operatorname{sign} \bar{G}_k \end{cases} \quad \text{when} \quad \begin{cases} k \in K^* \\ k \notin K^* \end{cases} \tag{3.4.2}$$

is, plainly, maximal, provided that $G \in \mathfrak{g}$; and these are all maximal G. Theorem 2 then holds and *all* minimal x^* satisfy (3.4.1). However, in general, such sequences G will not belong to $\mathfrak{g}$.

On the other hand, if $\mathfrak{g} = \mathfrak{l}$ [in which case the (minimal) kernel x of (3.1.1) is unique], then all the above G exist in $\mathfrak{l}$.

Theorem 3. *In order that the continuous linear functional b on $\mathfrak{l}$ should have a maximal sequence $A \in \mathfrak{l}$ it is necessary and sufficient that the kernel x of b should satisfy*

$$\sup|x_k| = \max|x_k| = \|b\|. \tag{3.4.3}$$

If $\tilde{K}$ is the non-empty index set $[k:|x_k| = \|b\|]$, then the maximal sequences A are exactly those, with $\|A\| = 1$, that satisfy

$$A_k = \begin{cases} 0 \\ \text{or} \quad x_k = \|b\| \operatorname{sign} \bar{A}_k \end{cases} \quad \text{when} \quad \begin{cases} k \in \tilde{K} \\ k \notin \tilde{K} \end{cases}. \tag{3.4.4}$$

3.5. If the $a_i = \{a_{ik}\}$, $1 \le i \le n$, $k \ge 0$, belong to $\mathfrak{A} = \mathfrak{l}$, that is, if they satisfy (0.1.2) for $p = 1$, then, by (3.1.1), the system (1.1.1) is equivalent to the system (0.1.1). Theorem 2, applied to the subspace $\mathfrak{P}_n$ of $\mathfrak{l}$ spanned by the a_i then yields Theorem B of the introduction.

The space $\mathfrak{l}$ is not reflexive. Hence for an infinite system (0.4.1) of linear equations, even if they have solutions, an "index theorem" will, in general, only hold if the corresponding c.l.f. $\tilde{b}$ has a maximal point in $\mathfrak{P}$. This, for instance, will be the case when $\mathfrak{P}$ itself is reflexive. In any case, the condition (3.4.1) for every minimal solution is a necessary condition for the existence of such a maximal point. If $\mathfrak{P} = \mathfrak{l}$, it is also sufficient.

As an example, where no index theorem applies, consider the infinite system

$$\begin{cases} 0 \cdot x_0 + \frac{1}{2}\lambda_1(x_1 + x_2) + \quad 0 \cdot (x_3 + x_4) + \quad 0 \cdot (x_5 + x_6) + \cdots = 1 \\ 0 \cdot x_0 + \quad 0 \cdot (x_1 + x_2) + \frac{1}{2}\lambda_2(x_3 + x_4) + \quad 0 \cdot (x_5 + x_6) + \cdots = 1 \\ 0 \cdot x_0 + \quad 0 \cdot (x_1 + x_2) + \quad 0 \cdot (x_3 + x_4) + \frac{1}{2}\lambda_3(x_5 + x_6) + \cdots = 1, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{cases} \tag{3.5.1}$$

where $\lambda_k > 1$, $k \ge 1$, and $\lambda_k \downarrow 1$. There exist solutions, for instance, $\tilde{x}_{2k} = \tilde{x}_{2k-1} = 1/\lambda_k$, $k \ge 1$, and $\tilde{x}_0$ arbitrary. Also, clearly, at least one of the two unknowns x_{2k-1}, x_{2k} must be not less than $1/\lambda_k$. Hence $\sup|x_k^*| = 1$ for any minimal solution (for instance, for the above $\tilde{x}$ if $|\tilde{x}_0| \le 1$). This allows for some

variation in each couple x_{2k-1}^*, x_{2k}^* while x_0^* is only restricted by $|x_0^*| \leq 1$. Hence there is no index K for which x_K^* must be the same for every minimal solution. Also there exist some minimal solutions for which $|x_k^*| < 1$ for all k.

4. The space c.

4.1. We consider the Banach space c of all complex (real) sequences $a = \{a_k\}$, $k \geq 1$, for which $a_k \to a_0$. The norm is $\|a\| = \sup |a_k|$. Every c.l.f. b on c can be represented in the form[10])

$$b(a) = \sum_{k=0}^{\infty} a_k x_k \tag{4.1.1}$$

where

$$\|b\| = \sum_{k=0}^{\infty} |x_k| = \|x\|_1. \tag{4.1.2}$$

The *kernel* $x = \{x_k\} \in \mathfrak{l}$ and is unique.

4.2. Let $\mathfrak{g}$ be a linear subspace of c. If $\tilde{b}$ is a c.l.f. defined on $\mathfrak{g}$, with norm $\|\tilde{b}\|_*$ on $\mathfrak{g}$, then, by the Hahn-Banach theorem, there exists a representation

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k x_k^* \qquad (g = \{g_k\} \in \mathfrak{g}) \tag{4.2.1}$$

where

$$\|\tilde{b}\|_* = \sum_{k=0}^{\infty} |x_k^*| = \|x^*\|_1. \tag{4.2.2}$$

For any other representation of $\tilde{b}$ with kernel x we have $\|x\|_1 \geq \|\tilde{b}\|_*$ so that again x^* is a ***minimal*** kernel for $\tilde{b}$. The conjugate space $\mathfrak{l}$ of c is not strictly convex. Hence, in general, we shall have an infinity of minimal kernels x^* for $\tilde{b}$. They form a convex set lying on the sphere $\|x\|_1 = \|\tilde{b}\|_*$ in $\mathfrak{l}$.

4.3. A sequence $G = \{G_k\} \in \mathfrak{g}$ is ***maximal*** for $\tilde{b}$ if

$$\text{(i)} \quad \|G\| = 1, \qquad \text{(ii)} \quad \tilde{b}(G) = \|\tilde{b}\|_*. \tag{4.3.1}$$

If $\|\tilde{b}\|_* = 0$, then all $G \in \mathfrak{g}$ with $\|G\| = 1$ are maximal. If $\|\tilde{b}\|_* > 0$, there may be no maximal sequences, even when $\mathfrak{g}$ is closed [since c is not reflexive]; and there may be many [since c is not strictly convex]. They form a convex set lying on the unit sphere $\|g\| = 1$ of $\mathfrak{g}$.

We now *assume* that $\tilde{b}$ has a maximal G. Then

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x_k^* = \|\tilde{b}\|_* = \sum_{k=0}^{\infty} |x_k^*|. \tag{4.3.2}$$

This implies

$$x_k^* = \begin{cases} 0 \\ 0 \end{cases} \quad \text{or} \quad \operatorname{sign} x_k^* = \operatorname{sign} \overline{G}_k \quad \text{whenever} \quad \begin{cases} |G_k| = 1 \\ |G_k| < 1 \end{cases}. \tag{4.3.3}$$

Hence either $x_k^* = 0$ eventually, or $\operatorname{sign} x_k^* \to \operatorname{sign} \overline{G}_0$ as $k \to \infty$ for the $x_k^* \neq 0$, where $G_0 = \lim G_k$: the sequence $x^* = \{x_k^*\}$ is directed. We also note that,

[10]) [1], p. 67 (for the real space c).

when $\|\tilde{b}\|_* > 0$, we must have at least one $x_k^* \neq 0$ and hence the corresponding $|G_k| = 1$. Also (4.3.3) must hold for every pair of a maximal G and a minimal x^*. Hence all minimal x^* are equally directed.

Theorem 4. *Suppose that* $\mathfrak{g}$ *is a subspace of* $\mathfrak{c}$ *and that the continuous linear functional* $\tilde{b}$ *on* $\mathfrak{g}$ *has a maximal sequence* G. *Then all minimal kernels* x^* *of* $\tilde{b}$ *are equally directed.*

Remarks. (i) If $\mathfrak{g} < \mathfrak{c}_0$ where $\mathfrak{c}_0$ is the subspace of all sequences g with $g_k \to 0$, then $G_0 = 0$. Hence there exists a K such that for all (equally directed) minimal kernels x^* of $\tilde{b}$

$$x_k^* = 0 \quad \text{for} \quad k > K. \tag{4.3.4}$$

(ii) In the case of the space $\mathfrak{c}_r$ of real convergent sequences, $|G_0| < 1$ or $G_0 = \pm 1$. It follows that there exists a K such that, for $k > K$, all minimal solutions x^* of $\tilde{b}$ have constant sign (non-negative, or non-positive).

4.4. Suppose that *some* minimal kernel x^{**} of $\tilde{b}$ is directed. Then any sequence G satisfying $\|G\| = 1$ and

$$|G_k| = 1, \quad \operatorname{sign} \overline{G}_k = \operatorname{sign} x_k^{**} \quad \text{whenever} \quad x_k^{**} \neq 0, \tag{4.4.1}$$

is, plainly, maximal for $\tilde{b}$ provided that $G \in \mathfrak{g}$. Theorem 4 then holds, and *all* x_k^* are equally directed and satisfy the analogue of (4.4.1). However, such $G \in \mathfrak{g}$ will not always exist. On the other hand, if $\mathfrak{g} = \mathfrak{c}$ (in which case the (minimal) kernel x of (4.1.1) is unique), then all these G are in $\mathfrak{c}$.

Theorem 5. *In order that the continuous linear functional* b *on* $\mathfrak{c}$ *should have a maximal sequence* $A \in \mathfrak{c}$ *it is necessary and sufficient that the kernel* x *of* b *should be directed.*

The maximal sequences A *are then exactly those that satisfy* $\|A\| = 1$ *and*

$$|A_k| = 1, \quad \operatorname{sign} \overline{A}_k = \operatorname{sign} x_k \quad \textit{whenever} \quad x_k \neq 0. \tag{4.4.2}$$

Remarks. (i) For the space $\mathfrak{c}_0$ the characteristic condition is that the kernel x of b should be "finite"; i.e., that $x_k = 0$ eventually. This kernel (normalized by $x_0 \neq 0$) is unique.

(ii) For the real space $\mathfrak{c}_r$, the condition is that the kernel x of b should be of constant sign eventually.

4.5. If the $a_i = \{a_{ik}\}$ in (0.1.1) belong to $\mathfrak{c}$, then the application of Theorem 4 to the subspace $\mathfrak{P}_n$ of $\mathfrak{c}$ spanned by the a_i yields Theorem C of the introduction and its corollaries Theorems C_r and C_0.

The space $\mathfrak{c}$ is not reflexive. Hence for an infinite system (0.4.1) of linear equations, even if they have solutions, a "direction theorem' will usually only hold if the corresponding c.l.f. $\tilde{b}$ has a maximal sequence in $\mathfrak{P}$. For the existence of a maximal sequence it is necessary that all minimal solutions should be equally directed. If $\mathfrak{P} = \mathfrak{c}$, this is also sufficient.

The system of equations (in c_0)

$$(4.5.1)\qquad \begin{cases} 1\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4+\cdots=1\\ 0\cdot x_1-4\cdot x_2+0\cdot x_3+0\cdot x_4+\cdots=1\\ 0\cdot x_1+0\cdot x_2+9\cdot x_3+0\cdot x_4+\cdots=1\\ \cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

has the unique solution $x=\left\{(-1)^{k-1}\frac{1}{k^2}\right\}\in l$. This solution is neither finite nor directed.

4.6. The space c is a subspace of the space l^∞ of all bounded sequences $a=\{a_k\}$, $k\geq 0$, with norm $\|a\|=\sup|a_k|$. The difficulty in dealing with l^∞ lies in the intricate form of its continuous linear functionals. The situation here is quite different from that in c. *Thus the system* (0.1.1) *with the* $a_i=\{a_{ik}\}\in l^\infty$ *need have no minmal solutions in* l. Consider, for instance, the one equation

$$(4.6.1)\qquad \sum_{k=0}^{\infty} a_k x_k=1,$$

where $0<a_k<1$ and $\sup a_k=1$. Clearly, for any solution $x=\{x_k\}\in l$ of (4.6.1) we must have $\|x\|=\sum_0^\infty |x_k|>1$. On the other hand, if $\xi_k^{(p)}=1/a_p$ when $k=p$, and $\xi_k^{(p)}=0$ otherwise then $\xi^{(p)}=\{\xi_k^{(p)}\}$ is a solution with $\|\xi^{(p)}\|=1/a_p$. Chosing suitable values of p, we see that $\inf\|x\|=1$, the inferior being taken with respect to all solutions x. It follows that (4.6.1) has no minimal solutions. *Suppose*, however, that the system (0.1.1), with $a_i\in l^\infty$, has minimal solutions $x^*\in l$. If then x^* and x^{**} are two minimal solutions, it is easy to see that

$$(4.6.2)\qquad \arg x_k^*=\arg x_k^{**},$$

whenever both $x_k^*\neq 0$ and $x_k^{**}\neq 0$, so that the *property* (β) (in the introduction) *for equal directedness still holds for the minimal solutions in this case, provided that they exist.*

For, clearly, $\frac{1}{2}(x^*+x^{**})$ is also a minimal solution, and hence

$$(4.6.3)\qquad \sum_{k=0}^{\infty}\tfrac{1}{2}|x_k^*+x_k^{**}|=\tfrac{1}{2}\sum_{k=0}^{\infty}\left(|x_k^*|+|x_k^{**}|\right),$$

which implies $|x_k^*+x_k^{**}|=|x_k^*|+|x_k^{**}|$ for all k, or (4.6.2).

References.

[1] BANACH, S.: Théorie des opérations linéaires. Warszawa 1932. — [2] BOHNENBLUST, H. F., and A. SOBCZYK: Extensions of functionals on complex linear spaces. Bull. Amer. Math. Soc. **44**, 91—93 (1938). — [3] RIESZ, F.: Les systèmes d'équations linéaires à une infinité d'inconnues. Paris 1913. — [4] RIESZ, F., and B. SZ-NAGY: Leçons d'analyse fonctionelle. Budapest 1952. — [5] ROGOSINSKI, W. W., u. H. S. SHAPIRO: On certain extremum problems for analytic functions. Acta Math. **90**, 287—318 (1953).

King's College, University of Durham, Newcastle upon Tyne (England).

(Eingegangen am 7. Januar 1955.)

DINGHAS, A.
Math. Zeitschr. Bd. 63, S. 109—132 (1955)

Konvexitätseigenschaften von Mittelwerten harmonischer und verwandter Funktionen.

ISSAI SCHUR zum Gedächtnis.

Von

ALEXANDER DINGHAS.

§ 1. Einleitung.

1.1. Voraussetzungen. Im folgenden bedeutet H^n den Halbraum $x_1>0$ des euklidischen Raumes R^n, dargestellt durch die Kartesischen Koordinaten $x_1, \ldots, x_n$. Wir setzen ferner $x_1=r\cos\vartheta$ mit $x_1^2+\cdots+x_n^2=r^2$ $(r>0)$ und $-\pi<2\vartheta<\pi$.

Die Funktionenklasse $L_0=L_0(u)$ sei nun folgendermaßen definiert:

1. Die Funktion $u=u(P)$ ist in H^n eindeutig und genügt der (Lindelöfschen) Bedingung

$$\overline{\lim_{P\to Q}}\, u(P)\leq 0 \quad (P\in H^n) \tag{1.1.1}$$

für jeden endlichen Randpunkt Q von H^n.

2. In der Umgebung jedes Punktes $P\in H^n$, wo $u(P)$ endlich bleibt, ist sie entweder harmonisch oder eine zweimal stetig differenzierbare subharmonische Funktion. Das bedeutet, daß $u(P)$ in der Umgebung jedes Punktes $P\in H^n$, wo sie endlich bleibt, der Bedingung

$$\frac{\partial^2 u}{\partial x_1^2}+\cdots+\frac{\partial^2 u}{\partial x_n^2}\geq 0 \tag{1.1.2}$$

genügt.

Danach ist eine Funktion der Klasse L_0 stets eine zweimal differenzierbare subharmonische Funktion, welche der Lindelöfschen Bedingung (1.1.1) genügt.

Eine Erweiterung der Klasse L_0 und somit auch des Gültigkeitsbereiches der Ergebnisse dieser Arbeit erhält man durch Grenzprozesse. Im folgenden fassen wir alle subharmonischen Funktionen u, welche sich als Limes von nicht zunehmenden Folgen von subharmonischen Funktionen der Klasse L_0 darstellen lassen, in einer Klasse, der Klasse L_2 zusammen. Die Teilklasse von L_2, die monotone, gleichmäßig konvergente Folgen verwendet, soll mit L_1 bezeichnet werden.

Es gilt offenbar $L_0\subseteqq L_1\subseteqq L_2$. Ferner folgt aus der Definition der Klassen L_1 und L_2, daß Konvexitätssätze, die für die Klasse L_0 bewiesen werden, durch Grenzübergang ohne weiteres auf diese Klassen ausgedehnt werden können.

Es sei zunächst $u \in L_0$, und es bezeichne S_r die Halbkugel

$$x_1^2 + \cdots + x_n^2 = r^2 \quad (x_1 > 0,\ r > 0). \tag{1.1.3}$$

Setzt man

$$\psi(P) = \overset{+}{u}(P) : \cos\vartheta \tag{1.1.4}$$

mit

$$\overset{+}{u}(P) = \operatorname{Max}\{0, u(P)\}, \tag{1.1.5}$$

so wird sich zeigen, daß $\psi(P)$ auf S_r beschränkt ist.

Aus der allgemeinen Theorie der subharmonischen Funktionen weiß man jetzt, daß die Punktmenge $E\,(u(P) > 0,\ P \in H^n)$ (kurz E), falls sie nicht leer ausfällt, zusammenhängende Mengen enthält, die sich bis ins Unendliche erstrecken.

Man bezeichne jetzt mit $d\omega_n$ das Flächenelement der Einheitskugel S

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \tag{1.1.6}$$

und bilde die Mittelwerte

$$m_k(r) = \int_{S_r} \psi^{2k}(P) \cos^2\vartheta \, d\omega_n. \tag{1.1.7}$$

Wird hier das Integral im Lebesgueschen Sinne genommen, so sind diese Funktionen für alle Funktionen der Klasse L_2 endlich und bei gegebenem u entweder für alle r gleich Null, oder von einem r an positiv.

Zweck der vorliegenden Arbeit ist zunächst, einige Konvexitätseigenschaften der Mittelwerte $m_k(r)$ zu studieren.

1.2. Formulierung der Hauptsätze. Wachstums- und Konvexitätseigenschaften der Mittelwerte (1.1.7) sind, soweit mir bekannt, bisher nur für $k = 1/2$ und $k = 1$ untersucht worden[1]). Der Fall eines beliebigen $k \geqq 1$ ist bisher in der Literatur noch nicht behandelt worden.

Die wichtigsten Ergebnisse der Arbeit lassen sich, was die Konvexität anbetrifft, durch folgende Sätze zum Ausdruck bringen:

Satz 1. *Für jedes $u \in L_2$ sind die Funktionen $m_k(r)^{1:k} r^{n-2}$ $(k \geqq 1)$ konvexe Funktionen von r^n.*

Satz 2. *Für jedes $u \in L_2$ sind die Funktionen $m_k(r)^{1:2k} : r$ konvexe Funktionen von $1 : r^n$.*

Will man nun den Grenzübergang $k \to \infty$ durchführen, so muß man zwei Fälle unterscheiden:

1. Die Funktion u liegt in der Klasse L_0 oder L_1. Dann gilt

$$\lim_{k \to \infty} m_k(r)^{1:2k} = \operatorname*{Max}_{P \in S_r} \{\psi(P)\} = M_1(r).$$

[1]) Man vgl. z.B. AHLFORS, L.: On Phragmén-Lindelöf's Principle. Trans. Amer. Math. Soc. **41**, 1—8 (1937). Hier werden ausschließlich harmonische Funktionen zugrunde gelegt, doch lassen sich die Ergebnisse ohne weiteres auf alle Funktionen der Klasse L_2 übertragen. — DINGHAS, A.: Über das Phragmén-Lindelöfsche Prinzip und einige andere verwandte Sätze. Sitzungsber. Preuß. Akad. Wissensch., Berlin, Math.-Physik. Kl. **12**, 123—140 (1938). — TSUJI, M.: On Harmonic Function on a Half-plane. Jap. J. of Math. **15**, 277—285 (1939).

2. Es ist $u \in L_2$. In diesem Falle definieren wir die Größe $M_1(r)$ folgendermaßen:

Es sei $P \in S_r$ und es bezeichne $M^*(r)$ das wahre Maximum von $\overset{+}{u}$ auf S_r, d.h. die untere Grenze aller Zahlen A, für welche die Menge $E(u(P) \geq A, P \in S_r)$ auf S_r gemessen, den Lebesgueschen Inhalt Null hat. Dann wird $M_1(r) = M^*(r)$ gesetzt.

Durch diese Definition von $M_1(r)$ kann man nun folgende allgemeine Sätze beweisen:

Satz 3. *Die Funktion $M_1(r)^2 r^{n-2}$ ist eine konvexe Funktion von $1 : r^n$.*

Aus dem zuletzt formulierten Satz 3 folgt ohne Schwierigkeit, daß, falls u nicht durchweg ≤ 0 ausfällt, der Ausdruck $m_k(r)^{1:2k} : r$, insbesondere auch der Ausdruck $M_1(r) : r$ für $r \to 0$ gegen positive Zahlen, die auch unendlich sein können, konvergieren. Das würde bedeuten:

Für jedes $u \in L_1$ genügt nicht bloß u dem Maximumprinzip, sondern auch die Funktion $u(P) : x_1$. Insbesondere trifft dies für den Fall zu, wenn u in H^n regulär harmonisch ist und der Lindelöfschen Bedingung (1.1.1) genügt.

Die Beweise findet der Leser in § 5.

Die beiden letzten Paragraphen der Arbeit beschäftigen sich mit folgender Frage: Kann man durch Erweiterung der Klasse erreichen, daß die Spanne bei der Aussage im klassischen Phragmén-Lindelöfschen Satz: ($u \leq 0$ in H^n oder mindestens von der Ordnung Eins) verringert wird?

Es zeigt sich, daß diese Frage positiv beantwortet werden kann, und daß man durch Hinzuziehung einer Klasse superharmonischer Funktionen die vorhandene Lücke ausfüllen kann. Für die genaue Formulierung der Resultate möge auf die Entwicklungen von § 7 hingewiesen werden, besonders auf die allgemeine Ungleichung (7.2.7).

Der letzte Paragraph erhält allgemeine Konvexitätssätze über subharmonische und harmonische Funktionen in R^n, insbesondere eine Übertragung des Hadamardschen Dreikreisesatzes auf diese Klassen.

Das wichtigste Resultat kann hier folgendermaßen formuliert werden: Man definiere die Klassen L_0^*, L_1^*, L_2^* genau wie die Klassen L_0, L_1, L_2 mit dem Unterschied, daß an Stelle H^n der ganze Raum R^n genommen wird und die Lindelöfsche Bedingung (1.1.1) wegfällt. Ebenfalls definiere man $M(r)$ für die Klasse L_2^*, ähnlich wie $M_1(r)$ für die Klasse L_2. Setzt man dann

$$\mu_k(r) = \{\int \overset{+}{u}(P)^{2k} d\omega_n\}^{1:2k} \quad (k \geq 1), \tag{1.1.8}$$

so gelten folgende Sätze:

Satz 4. *Die Funktion $\mu_k(r)$ $(k \geq 1)$ ist eine konvexe Funktion von $1 : r^{n-2}$.*

Satz 5. *Für $n = 2$ ist $M(r)$ eine konvexe Funktion von $\log r$. Dagegen ist für $n > 2$ $\mu_k(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.*

Geht man hier mit $k \to \infty$, so erhält man den

Satz 6. *Ist $n = 2$, so ist $M(r)$ eine konvexe Funktion von $\log r$. Für $n > 2$ ist $M(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.* Dieser Satz, der für $n = 2$ mit

dem Hadamardschen Dreikreisesatz übereinstimmt, gibt für $n>2$ seine natürliche Verallgemeinerung.

Die Bemerkungen des letzten Abschnitts von § 7 bewegen sich in der Richtung, die Frage der Existenz regulärer Integrale partieller Differentialgleichungen zu beantworten. Ich habe mich hier nur mit einigen Bemerkungen begnügt, da, soviel ich übersehe, die Arbeiten von Wittich diese Frage soweit gebracht haben, daß ich hier wesentlich neue Resultate dem Umfang nach kaum hinzufügen könnte.

1.2. Methode. Das Verfahren, das zum Beweis der hier formulierten Sätze verwendet wird, geht ursprünglich auf eine Methode von CARLEMAN[2]) zurück, der sie in etwas veränderter Form für $n=2$ und $k=1$ entwickelte. Der Leser wird leicht merken, wie stark diese Methode modifiziert werden muß, damit sie die allgemeinen und tieferliegenden Resultate der vorigen Nummern liefert.

Die gesamte Funktionentheorie einer komplexen Veränderlichen stützt sich eigentlich auf die Greensche Transformationsformel (4.1.3). Ich hoffe, in dieser Arbeit gezeigt zu haben, wie weit man auch Ungleichungsprobleme mit dieser Formel direkt angreifen und beantworten kann.

Im ganzen genommen ist diese Arbeit im Laufe eines Fortgeschrittenen-Kursus über Funktionentheorie, den ich in der Columbia University in New York mit Mitteln der Ford Foundation im Herbst 1952 gehalten habe, entstanden. Anderweitige Ablenkungen erlaubten mir jedoch erst jetzt[3]), das druckfertige Manuskript fertigzustellen.

Der Umweg über die Klasse L_0 und die Bildung der Klassen L_1 und L_2 scheint der natürlichste Weg zur Gewinnung der hier aufgezählten Resultate zu sein. Denn soweit es sich um subharmonische Funktionen handelt, scheinen wohl die Klasse L_1, L_2 sowie auch die Klasse L_1^*, L_2^* die allgemeinsten subharmonischen Funktionen zu enthalten. Ich hoffe, später bei Gelegenheit einer Darstellung dieses Fragenkomplexes (wie heutzutage üblich geworden ist) in einer kleinen Monographie dem Ganzen eine breitere Basis zu geben.

Zum Schluß möchte ich noch Herrn Dr. JOACHIM NITSCHE und Herrn KARL-HEINZ DIENER, für Hilfe bei der Korrekturarbeit und einige kritische Bemerkungen danken.

---

[2]) CARLEMAN, T.: Sur une inégalité differentielle dans la théorie des fonctions analytiques. Compt. Rend. Acad. Sci. Paris **196**, 995—997 (1933). Man vgl. auch DINGHAS, A.: Bemerkung zu einer Differentialgleichung von Carleman. Math. Z. **41**, 713—716 (1936). Ebenfalls: Sur un théorème de Denjoy-Carleman-Ahlfors. Bull. Soc. Math. Fr. **64**, 212 (1936).

[3]) Eine Zusammenfassung der Ergebnisse dieser Arbeit, allerdings nur eines Teiles, gab ich zuerst in einer Comptes Rendus Note: Sur quelques théorèmes concernant la convexité des moyennes d'une classe des fonctions sousharmoniques. Compt. Rend. Acad. Sci. Paris **237**, 594—595 (1953). — Für verwandte Probleme, zu welchen die Carlemansche Methode führt, vgl. man auch die Dissertation von HUBER: Über Wachstumseigenschaften gewisser Klassen von subharmonischen Funktionen. Zürich 1952. Dort findet der Leser auch weitere Literaturangaben. In ähnlicher Richtung bewegt sich die Dissertation von HANS-HEINRICH KELLER: Über das Anwachsen von Potentialfunktionen im dreidimensionalen Raum (Zürich 1950). Allgemeine Sätze und eine ausführliche Angabe der Eigenschaften der subharmonischen Funktionen gibt das ausgezeichnete Buch von T. RADO, Subharmonic Functions. Berlin 1937.

§ 2. Hilfssätze. Anfang.

2.1. Das Greensche Lemma und der Gauss-Ostrogradskische Satz. Damit der Leser alles beisammen haben möge, geben wir hier noch kurz den Wortlaut der beiden Sätze, auf den sich die nachfolgenden Entwicklungen stützen. Dabei werden soviel Voraussetzungen gemacht, als es für die Anwendung notwendig ist.

Wir bezeichnen im folgenden als regulären Bereich, eine beschränkte Punktmenge von R^n, die sich als Summe von endlich vielen Körpern im Sinne von ERHARD SCHMIDT[4]) darstellen läßt. Damit der Leser nicht die ganze Theorie zu lesen braucht, soll hier nur erwähnt werden, daß der Durchschnitt H der Punktmenge $E\left(u \geq \frac{1}{m}\right)$ $(m = 1, 2, 3, \ldots)$ mit jeder Halbkugel $x_1^2 + \cdots + x_n^2 \leq r^2$ $(x_1 > 0)$ von H^n ein regulärer Bereich ist, sobald $u \in L_0$ gilt.

Es bezeichne nun G den offenen Kern eines regulären Bereiches und Γ seinen Rand. Wir setzen dann $\overline{G} = G + \Gamma$. Ferner bezeichne ν die Richtung der äußeren Normale von $\overline{G}$, $d\tau$ das Volumenelement von G und dO das Oberflächenelement von Γ. Ist $\Phi = \Phi(P)$ eine in $\overline{G}$ stetig differenzierbare Funktion der Koordinaten $x_1, \ldots, x_n$ von P, so gilt für jedes k $(k = 1, \ldots, n)$

$$\int_{\overline{G}} \frac{\partial \Phi}{\partial x_k} d\tau = \int_{\Gamma} \Phi \cos(\nu, x_k)\, dO. \tag{2.1.1}$$

Diese Formel drückt das klassische Greensche Lemma aus. Hat man n Funktionen Φ_k mit den erwähnten Eigenschaften, so erhält man durch Summation

$$\int_{\overline{G}} \left\{ \sum_1^n \frac{\partial \Phi_k}{\partial x_k} \right\} d\tau = \int_{\Gamma} (\Phi, \nu)\, dO, \tag{2.1.2}$$

wobei hier Φ den Vektor mit den Koordinaten Φ_k bedeutet und die Klammer das innere Produkt von Φ mit ν darstellt.

Diese Gleichung drückt den Gauß-Ostrogradskischen Satz aus.

Der Fall, der uns im folgenden interessieren wird, wird durch die Gleichungen $\Phi_k = v \frac{\partial u}{\partial x_k}$ gegeben, wobei u, v zweimal stetig differenzierbare Funktionen bedeuten.

2.2. Vorbereitende Tatsachen. Einem Gebiet T von R^n sei die Metrik

$$ds^2 = g_{ik}\, dx^i\, dx^k \quad [5]) \tag{2.2.1}$$

aufgeprägt.

Wir betrachten einen regulären Bereich $\overline{G} \subseteq T$ und setzen voraus, daß wir eine Funktion $v = v(P)$ mit folgenden Eigenschaften bestimmen können:

1. Die Funktion $v(P)$ ist zweimal stetig differenzierbar in G und stetig in $G + \Gamma$.

[4]) SCHMIDT, ERHARD: Über das isoperimetrische Problem im Raum von n Dimensionen. Math. Z. **44**, 689—788 (1939).

[5]) Doppelte Indizes ziehen, wie üblich, eine Summation von 1 bis n nach sich, unabhängig voneinander.

2. Es ist $v(P)>0$ für $P\in G$.

3. Die Funktion $v(P)$ genügt in G der Differentialgleichung.

$$\Delta v+\lambda_0 v=0 \quad (2.2.2)$$

[6]) mit einem $\lambda_0>0$. Dabei bedeutet allgemein Δu den Lamé-Beltramischen Differentialausdruck

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^i}\left(\sqrt{g}\,g^{ik}\frac{\partial u}{\partial x^k}\right) \qquad (g=|g_{ik}|>0). \quad (2.2.3)$$

Es bezeichne nun U die Klasse der Funktionen u mit folgenden Eigenschaften:

1. Jedes u ist in G zweimal stetig differenzierbar, stetig in $G+\Gamma$, und Null auf Γ.

2. Setzt man

$$\varphi(P)=u(P):v(P), \quad (2.2.4)$$

so gilt für jeden Randpunkt $Q\in\Gamma$

$$\overline{\lim_{P\to Q}}|\varphi(P)|<+\infty. \quad (2.2.5)$$

3. Die Größen $\left|\frac{\partial u}{\partial x^k}\right|$ $(k=1, 2, \ldots, n)$ sind gleichmäßig beschränkt in G.

Definiert man, wie üblich, $\nabla(u, v)$ durch die Gleichung

$$\nabla(u,v)=g^{ik}\frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^k}, \quad (2.2.6)$$

so gilt für jedes $u\in U$

$$\int_G \nabla(u,u)\,d\tau_n<+\infty. \quad (2.2.7)$$

Dabei bedeutet $d\tau_n$ das Volumenelement $\sqrt{g}\,dx^1\ldots dx^n$ von G. Das Integral links wird hier (wie auch im folgenden) als Limes von Integralen $\int_{\bar G_k}\nabla(u,u)\,d\tau_n$ mit $\bar G_k\subseteqq\bar G_{k+1}<G$ $(k=1, 2, \ldots)$ und $\lim_{k\to\infty}\bar G_k=G$ definiert.

Sind $u_1, u_2\in U$, so gilt nach der Greenschen Transformation

$$\int_G u_1\,\Delta u_2\,d\tau_n=-\int_G \nabla(u_1,u_2)\,d\tau_n. \quad (2.2.8)$$

Infolge der Ungleichung

$$\left(\int_G|\nabla(u_1,u_2)|\,d\tau_n\right)^2\leq\int_G\nabla(u_1,u_1)\,d\tau_n\int_G\nabla(u_2,u_2)\,d\tau_n,$$

sind beide Seiten dieser Gleichung endlich.

2.3. Der Haupthilfssatz. Die Verallgemeinerung der Carlemanschen Methode wird nun durch folgenden wichtigen Hilfssatz ermöglicht:

Hilfssatz 1. *Die Funktion $f(x)$ möge im Intervall $I-\infty<x<+\infty$ folgende Eigenschaften besitzen:*

[6]) Für die allgemeinen Begriffsbildungen vgl. man: COURANT u. HILBERT: Methoden der mathematischen Physik, Bd. 1. Berlin 1931.

1. In jedem Punkt $x \in I$ besitzt $f(x)$ eine stetige Ableitung $f'(x) > 0$.
2. Es ist $f(0) = 0$.
Dann gilt für jedes $u \in U$

$$\int_G f(\varphi)\, v\,(\Delta u + \lambda_0 u)\, d\tau_n \leq 0, \tag{2.3.1}$$

und das Gleichheitszeichen gilt dann und nur dann, wenn in G $u = Cv$ mit einem konstanten C gilt, und $v(P)$ auf Γ verschwindet.

Beweis. Es sei $\overline{G}' \subset G$. Man betrachte das Integral

$$J = \int_{\overline{G}'} v^2 V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n, \tag{2.3.2}$$

wegen $v > 0$ auf dem Rand Γ' von $\overline{G}'$ kann man den Integranden entwickeln und schreiben

$$J = \int V(u,u)\, d\tau_n - 2\int \frac{u}{v} V(u,v)\, d\tau_n + \int \frac{u^2}{v^2} V(v,v)\, d\tau_n.$$

die Integration ist hier über $\overline{G}'$ zu erstrecken.

Es bedeuten nun ν und dO die äußere Normale bzw. das Oberflächenelement von Γ'. Dann wird durch Anwendung des Greenschen Lemmas

$$-2\int \frac{u}{v} V(u,v)\, d\tau_n = \int u^2 \frac{\partial}{\partial x^i}\left(\sqrt{g}\, g^{ik} \frac{1}{v} \frac{\partial v}{\partial x^k}\right) d\tau_n + J_0,$$

wobei J_0 das Flächenintegral $-\int_{\Gamma'} \frac{u^2}{v} \frac{\partial v}{\partial \nu}\, dO$ darstellt. Schreibt man jetzt J_1 für das Integral rechts, so wird nach einfachen Rechnungen

$$J_1 = \int \frac{u^2}{v} V v\, d\tau_n - \int \frac{u^2}{v^2} \Delta(v,v)\, d\tau_n.$$

Faßt man diese Ergebnisse zusammen, so erhält man mit Rücksicht auf (2.2.2)

$$J = J_0 + \int_{\overline{G}'} \{V(u,u) - \lambda_0 u^2\}\, d\tau_n.$$

Man schöpfe nun G durch eine monotone Folge $(\overline{G}'_k)$ $(k = 1, 2, \ldots)$ aus und beachte, daß bei genügend großem k der Betrag von u auf $\Gamma'_k \leq \varepsilon$ ausfällt. Läßt man also $k \to \infty$ konvergieren, so konvergieren die entsprechenden Integrale J_0 gegen Null, und man erhält

$$\int_G \{V(u,u) - \lambda_0 u^2\}\, d\tau_n = \int_G v^2 V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n. \tag{2.3.3}$$

Es seien nun $u_1, u_2 \in U$. Man ersetze in dieser Formel u durch die lineare Kombination $\alpha_1 u_1 + \alpha_2 u_2$ und vergleiche auf beiden Seiten die Koeffizienten von $\alpha_1 \alpha_2$.

Man erhält dann

$$\int_G \{V(u_1,u_2) - \lambda_0 u_1 u_2\}\, d\tau_n = \int_G v^2 V\left(\frac{u_1}{v}, \frac{u_2}{v}\right) d\tau_n \tag{2.3.4}$$

Jetzt setze man hier

$$u_1 = u \quad (u \in U), \qquad u_2 = f(\varphi)\, v \tag{2.3.5}$$

und wende das Greensche Lemma in der Form

$$\int_G V(u_1, u_2)\, d\tau_n = -\int_G u_2\, \Delta u_1\, d\tau_n \tag{2.3.6}$$

an. Dann wird

$$\int_G f(\varphi)\, v(\Delta u + \lambda_0 u)\, d\tau_n = -\int_G f'(\varphi)\, v^2\, V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n, \tag{2.3.7}$$

womit auch der Hilfssatz bewiesen wurde. Ich möchte hier erwähnen, daß dieser Hilfssatz auch unter denjenigen schwächeren Voraussetzungen gilt, unter welchen bei Zugrundelegung des Lebesgueschen Integrals die Greenschen Umformungen gelten. Ich gehe jedoch darauf nicht weiter ein.

§ 3. Hilfssätze. Ende.

3.1. Spezialfälle. Im folgenden ist der Fall, wo $f(x) = x^{2k-1}$ $(k \geq 1)$ ist und T mit der Halbkugel

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \quad (\xi_1 = \cos\vartheta \geq 0) \tag{3.1.1}$$

zusammenfällt, von besonderem Interesse. Wir zeigen vorerst, daß für jede Funktion $u \in L_0$ der Quotient $\overset{+}{u}(P) : \cos\vartheta$ auf jedem S_r beschränkt bleibt.

Hilfssatz 2. *Die Funktion*

$$\psi(P) = \overset{+}{u}(P) : \cos\vartheta \quad (P \in S_r,\ u \in L_0), \tag{3.1.2}$$

ist in jedem Intervall $0 < r_1 \leq r \leq r_2 < +\infty$ *nach oben beschränkt.*

Beweis. Ist $P(x_1, \ldots, x_n) \in H^n$, so soll P' den Punkt von R^n mit den Koordinaten $-x_1, x_2, \ldots, x_n$ bezeichnen. Es sei ferner H_R die Halbkugel

$$x_1^2 + \cdots + x_n^2 < R^2 \quad (x_1 > 0) \tag{3.1.3}$$

und $P \in H_R$. Wir setzen $OP = r$, $\varrho(Q) = PQ$ und $\varrho'(Q) = P'Q$. Ferner soll Q' den Punkt der Strecke OQ mit $OQ' = 1$ angeben.

Die Funktion

$$V(P) = \frac{R^{n-2}}{\omega_n} \int_{S_R} \overset{+}{u}(Q)\, \{\varrho^{-n}(Q) - \varrho'^{-n}(Q)\}\, (R^2 - r^2)\, d\omega_n(Q'),$$

wobei ω_n die Gesamtoberfläche von S bedeutet, ist in H_R harmonisch[7]), nimmt auf S_R die Randwerte $\overset{+}{u}(Q)$ und verschwindet auf dem ebenen Randteil $x_1 = 0$ von H_R. Da $u(P)$ subharmonisch und $u \in L_0$ ist, so gilt in H_R

$$u(P) \leq V(P). \tag{3.1.4}$$

Nun ist, wenn $y_1, \ldots, y_n$ die Koordinaten von $Q \in S_R$ bezeichnen

$$0 \leq \varrho^{-n}(Q) - \varrho'^{-n}(Q) \leq 2n\, x_1\, y_1 : (R - r)^{n+1},$$

[7]) Man vgl. z. B. CARATHÉODORY, C.: On Dirichlets Problem. Amer. J. Math. **59**, 709 (1937). Dieser Hinweis gilt natürlich nur für die dort entwickelte Poissonsche Formel. Der Leser kann daraus (3.1.4) selbst ableiten.

und infolgedessen wird für $P \in S_r$, $R = 2r$

$$\mathring{u}(P) : \cos\vartheta \leq C(r), \tag{3.1.5}$$

wobei $C(r)$ endlich ist. Das beweist den Hilfssatz. Es sei nun $u \in L_0$ und $E(u > 0)$ nicht leer. Bei genügend großem m ist dann $E\left(u \geq \frac{1}{m}\right)$ ebenfalls nicht leer.

Schreibt man hier wieder u an Stelle $u - \frac{1}{m}$, so gilt der

Hilfssatz 3. *Für jedes $k \geq 1$ gilt*

$$\int_{E_r} \psi^{2k-2}(P)\,(\Delta u + (n-1)\,u)\,u\,d\omega_n \leq 0, \tag{3.1.6}$$

und das Gleichheitszeichen tritt nur dann ein, wenn in H^n $u = C x_1$ mit einem konstanten C ist.

Hier bedeutet Δu den Beltramischen Operator (2.2.3) geschrieben mit Hilfe der g_{ik} der Einheitskugel $\xi_1^2 + \cdots + \xi_n^2 = 1$ und E_r die Menge $E(u \geq 0) \cap S_r$.

Bei der Anwendung des Hilfssatzes 1 hat man zuerst zu berücksichtigen, daß $\cos\vartheta$ der Differentialgleichung (2.2.2) mit $\lambda_0 = n - 1$ genügt[8]), und daß in jedem Punkt von E_r $\cos\vartheta > 0$ ausfällt. Ferner ist zu berücksichtigen, daß $\psi^{2k-2}(P)\cos\vartheta\,(u \in L_0)$ ebenfalls der Klasse U angehört, da ja $\psi^{2k-1}\cos\vartheta = u\,\psi^{2k-2}$ ist, und der Exponent nicht negativ ausfällt.

Der Hilfssatz 3 gilt selbstverständlich für viel breitere Funktionenklassen, als die Klasse L_0, z.B. für die Klasse L_2. Er wird aber ausschließlich für diese Klasse gebraucht und ich halte es nicht für notwendig, auf Vertiefungen einzugehen. Lediglich für den Fall $k = 1$ möchte ich hier eine Bemerkung anschließen, und zwar um den Eindruck zu zerstreuen, daß die hier gemachten Voraussetzungen enger sind als diejenigen des klassischen Wirtingerschen Lemmas[9]) für den Fall $n = 2$.

Wir setzen voraus, u sei auf der abgeschlossenen Halbkugel S_1

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \qquad (\xi_1 \geq 0), \tag{3.1.7}$$

als eindeutige Funktion von $P(\xi_1, \ldots, \xi_n)$ definiert und setzen

$$\xi_k = \sin\varphi_k\,\Pi_{k-1} \quad (k = 1, 2, \ldots, n-1), \qquad \xi_n = \Pi_n, \tag{3.1.8}$$

mit $\Pi_0 = 1$ und $\Pi_k = \prod_1^k \cos\varphi_r$. Die Variationsintervalle der φ_r sind hier einleuchtend.

Wir setzen nun ferner voraus:

1. Die Funktion u ist in S_1 stetig und genügt der Bedingung $u(0, \xi_2, \ldots, \xi_n) = 0$.

2. Die Funktion u ist in φ_1 totalstetig in der Umgebung von $\xi_1 = 0$.

3. Formt man u durch (3.1.8) um, so existiert $V(u, u)$ fast überall und ist im Lebesgueschen Sinne integrierbar. Dann kann bewiesen werden:

---

[8]) Vgl. z.B. DINGHAS, A.: Zur Theorie der konvexen Körper im n-dimensionalen Raum. Abh. Akad. Wiss. Berlin **1939**, No. 4.

[9]) Vgl. z.B. HARDY-POLYA-LITTLEWOOD: Inequalities. Cambridge 1934.

Es gilt stets

$$\int_{S_1} \{V(u,u) - (n-1)\,u^2\}\,d\omega_n \geq 0, \tag{3.1.9}$$

und das Gleichheitszeichen tritt nur für den Fall $u = C \sin \varphi_1$ ein.

Es bedeute in der Tat S_ϑ die Kugelzone $\sin \varphi_1 \leq \sin \vartheta$ und es bezeichne Γ_ϑ den Rand $\varphi_1 = \vartheta$ von S_ϑ.

Dann gilt, wie man leicht sieht,

$$\int_{\Gamma_\vartheta} \frac{u^2}{\sin \vartheta}\,dO_{n-1} \leq K \int_{S_\vartheta} V(u,u)\,d\omega_n.$$

Da nun wegen 3. die rechte Seite ein $o(1)$ ist, so wird

$$\int_{S_1} \{V(u,u) - (n-1)\,u^2\}\,d\omega_n = \int_{S_1} v^2\,V\left(\frac{u}{v}, \frac{u}{v}\right) d\omega_n,$$

mit $v = \sin \varphi_1$. Das beweist (3.1.9).

§ 4. Verallgemeinerung der Carlemanschen Methode.

4.1. Eine allgemeine Differentialungleichung. Vom Standpunkt der Greenschen Transformation aus gesehen erscheint die Carlemansche Methode als direkte Anwendung einer allgemeinen Transformationsformel, die in Spezialfällen mehrfache Anwendung findet. Es sei G ein regulärer Bereich und es seien dort $U = U(x_1, \ldots, x_n)$, $V = V(x_1, \ldots, x_n)$ zwei zweimal stetig differenzierbare Funktionen der rechtwinkligen Koordinaten $x_1, \ldots, x_n$. Man setze

$$(\operatorname{grad} U, \operatorname{grad} V) = \sum_1^n \frac{\partial U}{\partial x_k} \frac{\partial V}{\partial x_k}, \tag{4.1.1}$$

und

$$\operatorname{div} \operatorname{grad} U = \sum_1^n \frac{\partial^2 U}{\partial x_k^2}. \tag{4.1.2}$$

Dann gilt nach dem in 2.1 entwickelten Gaußschen Satz

$$\int_\Gamma V \frac{\partial U}{\partial \nu}\,dO_{n-1} = \int_{G+\Gamma} \{V \operatorname{div} \operatorname{grad} U + (\operatorname{grad} U, \operatorname{grad} V)\}\,d\tau_n. \tag{4.1.3}$$

Hierbei bedeutet $d\tau_n$ das Volumenelement $dx_1 \ldots dx_n$ von G und dO_{n-1} das Flächenelement des Randes Γ von G.

Man betrachte nun eine Funktion der Klasse L_0 und setze voraus, daß es mindestens einen Punkt P von H^n gibt, in dem $u(P) > 0$ ist. Betrachtet man dann die Funktionen $u_m = u - \frac{1}{m}$ $(m = 1, 2, \ldots)$, so gehören diese ebenfalls der Klasse L_0 an und genügen von einem m_0 an der Bedingung $u_m(P) > 0$. Man wähle nun ein festes $m > m_0$, schreibe der Einfachheit halber u an Stelle

von u_m und wende (4.1.3) auf den Durchschnitt G_r von $E(u>0)$ mit der Halbkugel S_r

$$x_1^2+\cdots+x_r^2\leq r^2 \qquad (x_1\geq 0) \tag{4.1.4}$$

an, indem man

$$v=2k\,\psi^{2k-1}(P)\cos\vartheta \qquad (\psi=\overset{*}{u}:\cos\vartheta)$$

wählt.

Nun wird

$$(\operatorname{grad}u,\operatorname{grad}v)=u_r v_r+\tfrac{1}{r^2}V(u,v), \tag{4.1.5}$$

und

$$\operatorname{div}\operatorname{grad}u=\tfrac{1}{r^{n-1}}(r^{n-1}u_r)_r+\tfrac{1}{r^2}\Delta u, \tag{4.1.6}$$

wobei $V(u,v)$, Δu die Beltramischen Differentialausdrücke von 3.1 für die Einheitskugel (3.1.1) bedeuten. Beachtet man jetzt, daß v auf dem Randteil von Γ, der nicht auf S_r liegt, verschwindet, so erhält man mit $v=\cos\vartheta$

$$r^{n-1}\int\limits_{E_r} v\,\frac{\partial u}{\partial r}\,d\omega_n=\int\limits_{G_r} f(P)\,t^{n-1}\,dt\,d\omega_n, \tag{4.1.7}$$

wobei der Kürze wegen

$$f(P)=v\left\{\tfrac{1}{r^{n-1}}(r^{n-1}u_r)_r+\tfrac{1}{r^2}\Delta u\right\}+\left\{u_r v_r+\tfrac{1}{r^2}V(u,v)\right\}$$

gesetzt wurde.

Schreibt man allgemein

$$F(x)=\int\limits_{E_x} f(P)\,d\omega_n, \tag{4.1.8}$$

so erhält man

$$r^{n-1}\int\limits_{E_r} v\,\frac{\partial u}{\partial r}\,d\omega_n=\int\limits_0^r x^{n-1}F(x)\,dx. \tag{4.1.9}$$

Will man nun diese Gleichung differenzieren, so sieht man leicht ein, daß stets

$$\frac{d}{dr}\left\{r^{n-1}\int\limits_{E_r} v\,\frac{\partial u}{\partial r}\,d\omega_n\right\}=r^{n-1}F(r) \tag{4.1.10}$$

ist, und zwar für diejenigen r, für welche $\int d\omega_n$ erstreckt über die Menge $E(u(P)=0,\ P\in S_r)$ verschwindet. In allen übrigen Stellen ist die untere Derivierte der linken Seite von (4.1.5) nicht kleiner als $r^{n-1}\int f(P)\,d\omega_n$, wobei das Integral über diejenige Teilmenge Γ_r von S_r zu erstrecken ist, auf der $u(P)>0$ ist.

Setzt man jetzt

$$m_k=m_k(r)=\int\limits_{\Gamma_r}\psi^{2k}(P)\cos^2\vartheta\,d\omega_n \qquad (k\geq 1), \tag{4.1.11}$$

so ist zunächst

$$m_k' = 2k \int_{\Gamma_r} \psi^{2k-1}(P)\, u_r \cos\vartheta\, d\omega_n,$$

und

$$\int_{\Gamma_r} u_r v_r\, d\omega_n = 2k(2k-1) \int_{\Gamma_r} \psi^{2k-2}(P)\, u_r^2\, d\omega_n. \tag{4.1.12}$$

Versteht man also unter $(r^{n-1} m_k')'$ in den Stellen, wo diese Ableitung nicht existiert, die untere Derivierte, und beachtet man, daß u subharmonisch ist, so wird

$$\frac{1}{r^{n-1}} (r^{n-1} m_k')' \geq 2k(2k-1) \int_{\Gamma_r} \psi^{2k-2} u_r^2\, d\omega_n + \frac{1}{r^2} \int_{\Gamma_r} V(u, v)\, d\omega_n.$$

Mit Hilfe der Greenschen Transformation erhält man

$$-\int_{\Gamma_r} V(u, v)\, d\omega_n = \int_{\Gamma_r} v\, \Delta u\, d\omega_n = 2k \int_{\Gamma_r} \psi^{2k-1}(P)\, \Delta u \cos\vartheta\, d\omega_n.$$

Ferner gilt

$$m_k'^2 \leq 4k^2 m_k \int_{\Gamma_r} \psi^{2k-2}(P)\, u_r^2\, d\omega_n.$$

Faßt man alle diese Resultate zusammen, so erhält man

$$\frac{1}{r^{n-1}} (r^{n-1} m_k')' \geq \frac{2k-1}{2k} \frac{m_k'^2}{m_k} + \frac{2k(n-1)}{r^2} m_k. \tag{4.1.13}$$

Es darf hier wiederholt werden, daß u stets $u - \frac{1}{m}$ bedeutet, und daß unter Umständen links die untere Derivierte von $r^{n-1} m_k'$ gemeint ist. Wegen der Stetigkeit von m_k' ist dies jedoch für das gesteckte Ziel unerheblich.

Die Ungleichung (4.1.13) ist für die nachfolgenden Ausführungen grundlegend. Nimmt man $k=1$, $n=2$ an, so wird

$$\frac{1}{r} (r m_1')' \geq \frac{1}{2} \frac{m_1'^2}{m_1} + \frac{2m_1}{r^2}, \tag{4.1.14}$$

d. h. die zu dem hier betrachteten Fall der Halbebene Carlemansche Ungleichung.

4.2. Umformungen. Anfang. Ist die Funktion u ($u \in L_0$) in H^n nicht durchweg ≤ 0, so ist von einem r an $m_k > 0$. Wir schreiben (4.1.13) in der Form

$$\frac{m_k''}{m_k} + \frac{n-1}{r} \frac{m_k'}{m_k} - \frac{2k-1}{2k} \left(\frac{m_k'}{m_k}\right)^2 \geq \frac{2k(n-1)}{r^2}$$

und formen diese Ungleichung folgendermaßen um.

Man setze zunächst

$$\mu(r) = m_k^{1:k}. \tag{4.2.1}$$

Dann erhält man für μ die Differentialungleichung

$$2\left(\frac{\mu''}{\mu} + \frac{n-1}{r} \frac{\mu'}{\mu}\right) - \left(\frac{\mu'}{\mu}\right)^2 \geq \frac{4(n-1)}{r^2}. \tag{4.2.2}$$

Führt man noch die Veränderliche ξ durch die Gleichung $\xi = r^n$ ein, und setzt man noch

$$y(\xi) = \mu(r)\, r^{n-2} = m_k(r)^{1/k} r^{n-2}, \tag{4.2.3}$$

so wird, falls die Ableitungen von y nach ξ durch $\dot{y}, \ddot{y}, \ldots$ bezeichnet werden,

$$\frac{\mu'}{\mu} = n \frac{\dot{y}}{y} r^{n-1} - \frac{n-2}{r}$$

und

$$\mu' = n \dot{y} r - \frac{n-2}{r^{n-1}} y.$$

Daraus folgt

$$\mu'' = n^2 \ddot{y} r^n - n(n-3) \dot{y} + \frac{(n-1)(n-2)}{r^n} y,$$

und somitwird

$$\frac{\mu''}{\mu} + \frac{n-1}{r} \frac{\mu'}{\mu} = n^2 r^{2n-2} \frac{\ddot{y}}{y} + 2n r^{n-2} \frac{\dot{y}}{y}.$$

Durch Einsetzen in (4.2.2) erhält man dann für y die Differentialungleichung

$$\ddot{y} \geq \frac{1}{2} y \left(\frac{\dot{y}}{y} - \frac{1}{\xi} \right)^2,$$

woraus dann schließlich

$$\frac{d^2 y}{d\xi^2} \geq \frac{1}{2} y \left(\frac{d}{d\xi} \log \frac{y}{\xi} \right)^2 \tag{4.2.1}$$

folgt.

4.3. Umformungen. Ende. Eine letzte Umformung von (4.1.13) sei noch gegeben. Man setze $\mu = g^2$ und $q = g : r = \mu^{\frac{1}{2}} : r = m_k^{1:2k} : r$.

Dann liefert diese Substitution die Differentialungleichungen

$$\frac{g''}{g} + \frac{n-1}{r} \frac{g'}{g} \geq \frac{n-1}{r^2} \tag{4.3.1}$$

und

$$r q'' + (n+1) q' \geq 0. \tag{4.3.2}$$

Da diese letztere sich in der Form

$$(r^{n+1} q')' \geq 0 \tag{4.3.3}$$

schreiben läßt, so genügt q als Funktion von $\eta = 1 : r^n$ der Differentialungleichung

$$\frac{d^2 q}{d\eta^2} \geq 0. \tag{4.3.4}$$

Hiermit sind alle Hilfsmittel bereitgestellt worden, die zum Beweis der in der Einleitung formulierten Sätze benötigt werden.

§ 5. Anwendungen.

5.1. Beweis der Sätze 1 bis 3. Da Konvexitätseigenschaften durch Anwendung von Grenzprozessen erhalten bleiben, so genügt es, die zwei ersten Sätze von 1.2 zunächst für Funktionen von der Form $u - \frac{1}{m}$ $(m = 1, 2, \ldots)$

mit $u \in L_0$ zu beweisen. Läßt man dann $m \to \infty$ konvergieren, so sind zunächst die Sätze 1 und 2 für jedes $u \in L_0$ bewiesen. Ein weiterer Grenzübergang $k \to \infty$ bei festem $u \in L_0$ beweist den Satz 3 für die Klasse L_0.

Damit der Leser alles beisammen haben möge, soll hier der bekannte Schluß wiederholt werden:

Es sei $f(x)$ in $a \leq x \leq b$ eindeutig stetig und ≥ 0. Dann gilt

$$J = \lim_{k \to \infty} \left\{ \int_a^b f(x)^k \, dx \right\}^{1:k} = \operatorname*{Max}_{a \leq x \leq b} \{f(x)\} = M. \tag{5.1.1}$$

In der Tat gilt zunächst trivialerweise $J \leq M$. Ist nun $\varepsilon > 0$, so gibt es ein Intervall $J_\varepsilon \langle a, b \rangle$ von der Länge $c J_\varepsilon > 0$, auf dem $f(x) \geq M - \varepsilon$ gilt. Daraus folgt $J \geq \lim_{k \to \infty} (c J_\varepsilon)^{1:k} (M - \varepsilon)$ also $J \geq M - \varepsilon$.

Wie verändert sich dieser Satz, wenn man die Stetigkeit aufhebt und dafür lediglich Meßbarkeit verlangt?

Wir zeigen:

Es sei M_0 das wahre Maximum von $f(x)$, d.h. die untere Grenze der Zahlen A, für welche die Menge $E(f(x) \geq A)$ das Lebesgue-Maß Null hat.

Dann gilt

$$\lim_{k \to \infty} \left\{ \int_a^b f(x)^k \, dx \right\}^{1:k} = M_0. \tag{5.1.2}$$

Es sei $\varepsilon > 0$. Man setze $f_\varepsilon(x) = f(x)$ für alle x, wo $f(x) \leq M_0 + \varepsilon$ ist und sonst $f_\varepsilon(x) = 0$.

Dann ist wegen

$$\int_a^b f^k(x) \, dx = \int_a^b f_\varepsilon^k(x) \, dx,$$

zunächst $J \leq M_0 + \varepsilon$. Andererseits ist

$$\int_a^b f^k(x) \, dx \geq (M_0 - \varepsilon)^k m_\varepsilon$$

mit einem $m_\varepsilon > 0$. Daraus folgt $J \geq M_0 - \varepsilon$, d.h. die Behauptung.

Ist nun $u \in L_0$, so kann ohne weiteres als bewiesen angesehen werden, daß $m_k(r) \, r^{n-2}$ als Grenzwert von monoton wachsenden konvexen Funktionen, d.h. der mit Hilfe der Funktionen $u - \frac{1}{m}$ gebildeten Ausdrücke (4.1.11), ebenfalls konvex ist. Darunter wird allgemein die Ungleichung

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2) \tag{5.1.1}$$

für jedes Wertetripel $x_1 < x < x_2$ verstanden.

Ist $u \in L_2$, so gibt es eine monoton abnehmende Folge (u_m) $(m = 1, 2, \ldots)$ von zweimal stetig differenzierbaren subharmonischen Funktionen der Klasse

L_0 mit $\lim u_m(P)=u(P)$ $(P\in H^n)$. Faßt man dann die Ausdrücke (4.1.11) als Lebesguesche Integrale auf, so erhält man durch Grenzübergang $m\to\infty$ den Beweis der Sätze 1 und 2 für die Klasse L_1. Faßt man schließlich für jedes $u\in L_2$ den Ausdruck $M_1(r)$ im Sinne der Gl. (5.1.2) auf, so genügt der Grenzübergang $k\to\infty$, um auch den Satz 3 für die Klasse L_2 zu beweisen.

Die Eigenschaft der Konvexität derAusdrücke $m_k^{1:k}\,r^{n-2}$ bzw. $m_k^{1:2k}:r$ als Funktion von r^n bzw. $1:r^n$ zieht die Existenz der Grenzwerte

$$a_k=\lim_{r\to\infty} m_k^{1:2k}(r):r \tag{5.1.2}$$

und

$$b_k=\lim_{r\to 0} m_k^{1:2k}(r)\,r^{n-1} \tag{5.1.3}$$

nach sich.

Insbesondere existiert Grenzwert

$$a=\lim_{r\to\infty} M_1(r):r. \tag{5.1.4}$$

Wie nochmals darauf hingewiesen werden darf, ist $M_1(r)$ nicht durch die Maximumeigenschaft definiert, sondern durch die Bedingung, daß die Teilmenge von S_r für die $u\geq M_1(r)+\varepsilon$ ist, das Maß Null (auf S_r!) hat, während dies für $M_1(r)-\varepsilon$ nicht zutrifft.

5.2. Scharfe Fassung des Phragmén-Lindelöfschen Prinzips. Ist $u\in L_1$, so kann man das klassische Phragmén-Lindelöfsche Prinzip schärfer als bisher fassen.

Es bedeute wieder H_r die Halbkugel

$$x_1^2+\cdots+x_n^2<r^2 \qquad (x_1>0,\ r>0) \tag{5.2.1}$$

und es sei $P(x_1,\ldots,x_n)\in H_r$.

Dann kann man beweisen:

Entweder ist die Funktion

$$m(P)=\overset{+}{u}(P):x \tag{5.2.2}$$

in H_r identisch Null oder sie ist stetig und nimmt ihr Maximum auf dem Rande S_r von H_r an.

Der Beweis des nichttrivialen Teiles dieses Satzes folgt aus der Tatsache, daß für je zwei Punkte $P(x_k)\in S_r$ und $P'(x_k')\in S_{r'}$ $(r<r')$

$$\frac{\overset{+}{u}(P)}{x_1}\leq \operatorname*{Max}_{Q\in S_r}\frac{\overset{+}{u}(Q)}{x_1}\leq \operatorname*{Max}_{Q'\in S_{r'}}\frac{\overset{+}{u}(P')}{x_1'}$$

gilt.

Für den Fall $n=2$ und einer in der rechten Halbebene $x>0$ regulären analytischen Funktion $f(z)$ der komplexen Veränderlichen $z=x+iy$ liefert dieser Satz den klassischen Satz von Phragmén und Lindelöf in folgender scharfer Form:

Ist $f(z)$ in $x>0$ regulär analytisch, und gilt für jeden endlichen Punkt ζ der imaginären Achse

$$\lim_{z\to\zeta}|f(z)|\leq 1 \tag{5.2.4}$$

so gilt entweder $|f(z)|\leq 1$ in $x>0$ oder der Ausdruck $\overset{*}{\log}|f(z)|:x$, der dort nicht identisch verschwindet, genügt für Halbkreisgebiete dem Maximumprinzip.

Diese Aussage kann auch so formuliert werden:

Die Funktion

$$\varphi(r)=\operatorname*{Max}_{z\in S_r}\{\overset{*}{\log}|f(z)|:x\} \tag{5.2.6}$$

ist eine nicht abnehmende Funktion von r und konvergiert infolgedessen für $r\to\infty$ gegen eine (endliche oder unendliche) Zahl $a>0$.

§6. Wachstum spezieller Funktionenklassen.

6.1. Erweiterung der Klasse L_0. Genügt eine Funktion $u=u(P)$ in H^n den zwei Bedingungen von 1.1, so kann sie dort nach dem eben bewiesenen Satz von Phragmén und Lindelöf entweder beschränkt sein oder mindestens vom Normaltypus von der Ordnung Eins. Durch eine Lockerung der Bedingungen soll es nun möglich werden, die Spanne dieser Alternativen zu verringern.

Eine in H^n reelle eindeutige Funktion $u=u(P)$ soll dort von der Klasse $\overline{L}_0$ heißen, wenn sie folgenden Bedingungen genügt.

1. Die Menge $E(u>0)$ ist nicht leer.

2. In der Umgebung jedes Punktes $P\in H^n$ ist u zweimal stetig differenzierbar.

3. In jedem Punkt der Menge $E(u\geq 0)$ genügt u der Differentialungleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2}\geq \frac{c}{r^2}u, \tag{6.1.1}$$

wobei $c=c(P)$ eine in H^n eindeutige stetige Funktion, ist, die der Ungleichung

$$c(P)+n-1\geq 0 \tag{6.1.2}$$

genügt.

4. Die Funktion u genügt der (Lindelöfschen) Bedingung

$$\overline{\lim_{P\to Q}}\,u(P)\leq 0 \qquad \big(P\in H^n,\ Q\in E(x=0)\big). \tag{6.1.3}$$

Infolge der Bedingung 3 braucht hier u keine subharmonische Funktion zu sein. Es wird jedoch gezeigt, daß sie durch die Einschränkung (6.1.2) Eigenschaften beibehält, die sie sehr nahe an die Funktionen der Klasse L_0 bringen.

6.2. Ungleichungen. Setzt man in (4.1.3) $U=V=u$ $(u\in L_0)$, so kann man ohne weiteres zeigen, daß die Punktmenge $E(u\geq 0)$ keine Inselgebiete enthalten kann, auf deren Rand Γ durchweg $u=0$ gilt. Ist nämlich G ein

solches im Endlichen liegendes Gebiet, so wird nach (4.1.3), wenn ν die Richtung der äußeren Normale bedeutet,

$$\int_{\Gamma} u \frac{\partial u}{\partial \nu} dO_{n-1} = \int_{G+\Gamma} \{u \operatorname{div} \operatorname{grad} u + (\operatorname{grad} u)^2\} d\tau_n.$$

Führt man rechts Polarkoordinaten ein, so wird die geschweifte Klammer mit Rücksicht auf (6.1.1) $\geq u_r^2 + \frac{1}{r^2}\{V(u,u) + c\,u^2\}$. Verwendet man jetzt das Wirtingersche Lemma (3.2.3) ($k=1$!) und (2.3.5) auf jedem Schnitt E_r von G mit der Halbkugel S_r, wonach

$$\int_{E_r} V(u,u)\, d\omega_n \geq (n-1) \int_{E_r} u^2\, d\omega_n \tag{6.2.1}$$

ist, so wird mit Rücksicht auf (6.1.2)

$$\int_{\Gamma} u \frac{\partial u}{\partial r} dO_{n-1} \geq \int_{G} u_t^2\, t^{n-1}\, d\omega_n\, dt. \tag{6.2.2}$$

Ist nun $u=0$ auf Γ, so müssen in G folgende Bedingungen erfüllt sein:

1. $$\int_{E_r} V(u,u)\, d\omega_n = (n-1) \int_{E_r} u^2\, d\omega_n$$

für alle r, die den Punkten von G entsprechen.

2. $$c(P) + n - 1 = 0$$

3. $$u_r = 0 \quad (P \in G)$$

4. $$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} = \frac{c}{r^2} u.$$

Wegen 3. und 2. schreibt sich diese Bedingung in der Form

$$u_{rr} + \frac{n-1}{r} u_r + \frac{1}{r^2} \Delta u = \frac{c}{r^2} u. \tag{6.2.3}$$

Dabei ist Δ, wie nochmals wiederholt werden darf, der zweite Beltramische Operator auf der Einheitskugel S.

Nun folgt aus 1. in Verbindung mit der Stetigkeit von u, daß E_r mit S_r übereinstimmen muß. Wegen 3. kann also u nur von der Form $C \cos \vartheta$ mit einem konstanten C sein. In diesem Falle ist dann (6.2.3) automatisch erfüllt.

Man bilde jetzt mit $v = u$ das Analogon von (4.1.7) und beachte, daß, falls dieser Ausnahmefall, der uns im folgenden nicht weiter interessieren wird, eintritt, von einem r an die rechte Seite der Gleichung

$$r^{n-1} \int_{E_r} u \frac{\partial u}{\partial r} d\omega_n = \int_{G_r} f(P)\, t^{n-1}\, dt\, d\omega_n \tag{6.2.4}$$

mit

$$f(P) = u\left\{\frac{1}{r^{n-1}} (r^{n-1} u_r)_r + \frac{1}{r^2} \Delta u\right\} + \left\{u_r^2 + \frac{1}{r^2} V(u,u)\right\}$$

positiv ist. Es genügt also hier

$$m(r) = \int_{S_r} \mathring{u}^2 \, d\omega_n \tag{6.2.5}$$

zu setzen, um nach dem Vorbild von 4.1 die Carlemansche Differentialgleichung zu bilden. Das soll nun in der nächsten Nummer geschehen.

6.3. Die Fundamentalungleichung. Man wiederhole hier nun den Gedankengang von 4.1 unter Berücksichtigung der veränderten Voraussetzungen über u, indem man hier wieder mit dem Gebiet G_r operiert.

Nimmt man r groß genug, damit $m > 0$ ausfällt, so wird

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geq \frac{1}{2} \frac{m'^2}{m} + \frac{2}{r^2} \int_{\Gamma_r} c\, u^2 \, d\omega_n + \frac{2}{r^2} \int_{\Gamma_r} \nabla(u, u) \, d\omega_n .$$

Durch Anwendung des Wirtingerschen Lemmas läßt sich diese Ungleichung auch in der Form schreiben

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geq \frac{1}{2} \frac{m'^2}{m} + \frac{2(n-1)}{r^2} m + \frac{2}{r^2} \int_{\Gamma_r} c\, u^2 \, d\omega_n .$$

Wie in dem Fall einer subharmonischen Funktion kann m'' unter Umständen die untere Derivierte von m' bedeuten. Das bleibt jedoch auf die Frage der Konvexität ohne Einfluß. Es sei nun

$$q = q(r) = \operatorname*{Min}_{P \in S_r} c(P). \tag{6.3.1}$$

Dann wird

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geq \frac{1}{2} \frac{m'^2}{m} + \frac{2(q+n-1)}{r^2} m . \tag{6.3.2}$$

Diese Ungleichung, die von einem r an gilt, schränkt das Wachstum der Funktionen der Klasse $\overline{L}_0$ ein. Der nächste Paragraph soll der Entwicklung der Konsequenzen von (6.3.2) gewidmet werden.

§7. Konvexitätssätze und Phragmén-Lindelöfsches Prinzip für die Klasse L_2. Der Hadamardsche Dreikugelsatz.

7.1. Die Klasse $\overline{L}_2$ und die Konvexitätssätze. Wie beim Fall der subharmonischen definiert man hier die Klasse $\overline{L}_2$ durch Hinzufügung der Grenzfunktionen aller monoton abnehmenden, konvergenten Folgen von Funktionen der Klasse $\overline{L}_0$. Die Klasse, welche aus allen gleichmäßig konvergenten monoton abnehmenden Folgen von Funktionen der Klasse $\overline{L}_0$ besteht, wird entsprechend mit $\overline{L}_1$ bezeichnet. Jedes $u \in \overline{L}_1$ ist dann in H^n stetig.

Infolge der Veränderlichkeit von $q(r)$ können Konvexitätssätze für die Klassen $\overline{L}_0$, $\overline{L}_1$ und $\overline{L}_2$ nur für feste Intervalle $J(r_1 < r < r_2)$ gegeben werden.

Es sei $q(J)$ durch die Gleichung

$$q(J) = q = \inf_{r \in J} q(r) \tag{7.1.1}$$

definiert und $\lambda = \lambda(J)$ durch die Gleichung

$$\lambda(\lambda + n - 1) = q + n - 1 \tag{7.1.2}$$

bestimmt.

Setzt man dann $\xi = r^{2\lambda+n-2}$ und

$$y(\xi) = m(r)\, r^{n-2} \tag{7.1.3}$$

so transformiert sich (6.3.2) wegen

$$\frac{m'}{m} = (2\lambda + n - 2)\,\frac{\dot{y}}{y}\, r^{2\lambda+n-3} - \frac{n-2}{r}$$

und

$$\frac{m''}{m} + \frac{n-1}{r}\,\frac{m'}{m} = (2\lambda + n - 2)^2\,\frac{\ddot{y}}{y}\, r^{2(2\lambda+n-3)} + 2\lambda(2\lambda + n - 2)\,\frac{\dot{y}}{y}\, r^{2\lambda+n-4}$$

in die Ungleichung

$$\frac{d^2 y}{d\xi^2} \geq \frac{1}{2}\, y \left(\frac{d}{d\xi} \log \frac{y}{\xi}\right)^2. \tag{7.1.4}$$

Es gilt also der Satz:

Ist $r \in J$ und wird λ durch (7.1.2) *definiert, so ist dort $m(r)\, r^{n-2}$ eine konvexe Funktion von $r^{n+2\lambda-2}$.*

Schreibt man wieder g für $m(r)^{\frac{1}{2}} : r^{\lambda}$, so genügt g der Differentialungleichung

$$g'' + \frac{n + 2\lambda - 1}{r}\, g' \geq 0, \tag{7.1-5}$$

aus der, ähnlich wie vorhin, der Satz folgt:

Die Funktion $m^{\frac{1}{2}} : r^{n+2\lambda-2}$ ist in J eine konvexe Funktion von $1 : r^{n+2\lambda-2}$.

Schreibt man die Konvexitätseigenschaften in der Form der Ungleichung (5.1.1), so kann man durch Grenzübergang die beiden eben bewiesenen Sätze für alle Funktionen der Klasse $\overline{L}_2$ beweisen. Wir gehen darauf nicht näher ein.

Die beiden Sätze sind insofern genau, als sie auch den linearen Fall einschließen. Das zeigen die Beispiele $u r^{\alpha} = \cos\vartheta$ $(\alpha > 0)$.

7.2. Nochmals das Phragmén-Lindelöfsche Prinzip. Ist $u \in L_0$, so definiere man die Folge

$$q_m = \inf_{r \geq m} \{q(r)\} \qquad (m = 1, 2, \ldots) \tag{7.2.1}$$

und setze

$$q_0 = \lim_{m \to \infty} q_m. \tag{7.2.2}$$

Schreibt man dann λ_m für $\lambda(q_m)$ und λ_0 für $\lambda(q)$, so kann man zeigen:

1. Für jedes m existieren stets die Grenzwerte

$$\lim_{r \to \infty} m(r)^{\frac{1}{2}} : r^{q_m} = A_m \qquad (m = 1, 2, \ldots). \tag{7.2.3}$$

2. Für jedes m ist $A_m > 0$.

In der Tat existieren wegen der Konvexität von $m(r)\, r^{n-2}$, in Verbindung mit der Tatsache $m' > 0$, die nichtnegativen Grenzwerte

$$\lim_{r\to\infty} \frac{m(r)\, r^{n-2} - m(r_1)\, r_1^{n-2}}{r^{2\lambda_m+n-2} - r_1^{2\lambda_m+n-2}} = C_m \qquad (m = 1, 2, \ldots) \tag{7.2.4}$$

und infolgedessen auch (durch Festhalten von r_1) die Grenzwerte

$$A_m = \lim_{r\to\infty} m(r) : r^{2\lambda_m} = C_m. \tag{7.2.5}$$

Das beweist die Behauptung.

Die Folge (λ_m) ist ähnlich wie die Folge (q_m) nicht abnehmend und infolgedessen kann man im allgemeinen keine Aussage über das Verhalten des Ausdrucks $m(r) : r^{2\lambda_0}$, falls $r \to \infty$ konvergiert, machen. Die einzige allgemeine Aussage, die man hier machen kann, ist die, daß nämlich stets

$$\varliminf_{r\to\infty} \frac{\log m(r)^{\frac{1}{2}}}{\log r} \geq \lambda_0 \tag{7.2.6}$$

gilt, also auch

$$\varliminf_{r\to\infty} \frac{\log M(r)}{\log r} \geq \lambda_0. \tag{7.2.7}$$

Hierbei bedeutet, wie üblich, $M(r)$ das Maximum von $\overset{+}{u}(P)$ auf S_r. Die Zahl $\lambda_0 \left(\lambda_0 = \lim_{m\to\infty} \lambda(q_m)\right)$ kann hier alle nichtnegativen Werte annehmen und insofern bedeutet die Aussage (7.2.7) eine Vertiefung des Phragmén-Lindelöfschen Prinzips, das lediglich die Werte $\lambda_0 \geq 1$ zuläßt.

Wie die Funktionen

$$u(P) = \left(r^{\alpha} \prod_1^{\varrho} \log_k^{\alpha_k} r\right) \cos\vartheta \qquad (\alpha_k \geq 0),$$

zeigen, für die

$$c(P) = \alpha(\alpha + n - 2) - (n-1) + o(1)$$

gilt, und insofern in (7.2.7) das Gleichheitszeichen eintritt, ist kaum zu erwarten, wie im Falle der harmonischen Funktionen, eine Aussage über den Typus von u zu erhalten.

Im Falle $\lambda_0 > 0$ ist das Resultat (7.2.7) insofern bemerkenswert, als es zeigt, daß man durch Hinzunahme der Klasse der subharmonischen Funktionen mit $-(n-1) \leq c(P) \leq 0$ die Spanne im Phragmén-Lindelöfschen Prinzip (beschränkt oder mindestens von der Ordnung Eins) beliebig verringern kann.

7.3. Der Hadamardsche Dreikugelsatz in R^n. Die Klasse S_2 der subharmonischen Funktionen von R^n besteht bekanntlich aus allen zweimal stetig differenzierbaren subharmonischen Funktionen (Klasse S_0) von R^n und

allen nicht zunehmenden, konvergenten Folgen von S_0. Früheren Bezeichnungen entsprechend bezeichnen wir die Klasse aller stetigen subharmonischen Funktionen, die Limes von monoton abnehmenden, gleichmäßig konvergenten subharmonischen Funktionen der Klasse S_0 sind, mit S_1.

Nachfolgender Satz ist nicht tief und soll hier nur der Vollständigkeit wegen gebracht werden. Ist $u \in S_0$ und bildet man die Mittelwerte

$$\varrho_k(r) = \int_{S_r} u^+(P)^{2k}\, d\omega_n, \tag{7.3.1}$$

wobei S_r hier die volle Kugel $x_1^2 + \cdots + x_n^2 = r^2$ bedeutet, so kann man leicht die Differentialungleichung

$$\frac{1}{r^{n-1}}\,(r^{n-1}\,\varrho_k')' \geq \frac{2k-1}{2k}\,\frac{\varrho_k'^2}{\varrho_k}, \tag{7.3.2}$$

d.h.

$$\frac{\varrho_k''}{\varrho_k} + \frac{n-1}{r}\,\frac{\varrho_k'}{\varrho_k} - \frac{2k-1}{2k}\left(\frac{\varrho_k'}{\varrho_k}\right)^2 \geq 0 \tag{7.3.3}$$

gewinnen. Setzt man hier $\mu_k = \varrho_k^{1:2k}$, so wird nach leichten Umrechnungen

$$(r^{n-1}\mu_k')' \geq 0. \tag{7.3.4}$$

Ist nun $n = 2$, so drückt diese Ungleichung die Konvexität von μ_k als Funktion von $\log r$, und zwar (nach geeignetem Grenzübergang) für jedes $u \in S_0$. Drückt man wieder die Konvexität von μ_k in der Form (5.1.1) aus, und geht man mit $k \to \infty$, so erhält man die Behauptung für $M(r)$ für die Klassen S_0 und S_1. Die Klasse S_2 erledigt sich durch die schon vorhin gegebene Erweiterung der Definition von $M(r)$. Ist nun $n > 2$ und setzt man $y = \mu(r)\, r^{n-2}$, $\xi = r^{n-2}$, so erhält man aus (7.3.4)

$$\frac{y''}{y} - \frac{n-3}{r}\,\frac{y'}{y} \geq 0,$$

d.h. $\frac{d^2 y}{d\xi^2} \geq 0$. Das Übrige, die Konvergenz $k \to \infty$ und der Übergang zu den Klassen S_1, S_2 erledigt sich wie vorhin. Zusammenfassend können wir den Dreikugelsatz von Hadamard in folgender Form beweisen:

Es sei $n > 2$ und $u \in S_2$. Definiert man für jedes r die Größe $M(r)$ als die untere Grenze der Zahlen M, für die $E(u > M,\ P \in S_r)$ das Maß Null hat, so ist $M(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.

7.4. Nichtlineare Fälle. Um zu zeigen, daß die Klasse S_2 eine ziemlich abgeschlossene Klasse darstellt, will ich, ohne damit Vollständigkeit anzustreben, zeigen, daß unter bestimmten einfachen Wachstumsvoraussetzungen die Differentialgleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} = g\big(P, u(P)\big), \tag{7.4.1}$$

sobald g in u nicht linear ist, keine durchweg in R_n regulären Integrale besitzen kann. Der Fall einer Ungleichung erledigt sich dann analog.

Der Einfachheit halber setze ich $n=2$ voraus und zeige:

Gilt bei vorgegebenem $\varepsilon>0$ von einem $r=r(\varepsilon)$ an

$$g\left(P, \overset{+}{u}(P)\right) \geq \frac{C}{r^2} \overset{+}{u}^{1+4\varepsilon} \quad \left(C=C(\varepsilon)\right), \tag{7.4.2}$$

so kann (7.4.1) keine im Endlichen regulären Integrale haben.

Zunächst kann man ohne weiteres annehmen, daß es mindestens einen Punkt P_0 gibt, in dem $u(P_0)>0$ ist. Dann zeigt man, daß $E(u>0)$ keine im Endlichen liegenden Gebiete enthalten kann, auf deren Rand $u=0$ ist.

Man setze nun zunächst $u \in S_0$ voraus und verwende wieder den Ausdruck

$$m(r)=\frac{1}{2\pi}\int_0^{2\pi} \overset{+}{u}^2(P)\, d\varphi \quad (P \in |z|=r). \tag{7.4.3}$$

Setzt man hier $r \geq r(\varepsilon)$ voraus, so wird wegen

$$\frac{1}{2\pi}\int_0^{2\pi} \overset{+}{u}^{2+4\varepsilon}\, d\varphi \geq m^{1+2\varepsilon},$$

$$\left(\frac{m''}{m'}+\frac{1}{r}\right)^2 \geq 2\left(\frac{m''}{m'}+\frac{m'}{rm}\right)-\left(\frac{m'}{m}\right)^2 \geq \frac{C}{r^2} m^{2\varepsilon}. \tag{7.4.4}$$

Daß hier m von einem r an positiv ist, und daß links der eingeklammerte Ausdruck nicht negativ ist, sieht man aus den Entwicklungen von 4.1.

Aus (7.4.4) folgt nun für $r \geq r(\varepsilon)$

$$\frac{m''}{m'}+\frac{1}{r} \geq \frac{C_1}{r} m^{\varepsilon} \quad (C_1=C_1(\varepsilon)>0)$$

oder wenn $r=e^{\xi}$ gesetzt wird,

$$\frac{d}{d\xi}\log\frac{dm}{d\xi} \geq C_1 m^{\varepsilon}.$$

Das gibt die Ungleichung

$$\frac{d^2m}{d\xi^2} \geq C_2 \frac{dm^{1+\varepsilon}}{d\xi},$$

aus der dann durch Integration die Ungleichung

$$\int^{r} \frac{dm}{m^{1+\varepsilon}} \geq C_2 \int^{r} \frac{dr}{r}+C_3 \quad (C_2>0) \tag{7.4.5}$$

mit einem konstanten C_2, C_3 folgt. Das beweist die Behauptung. Der allgemeine Fall $u \in S_2$ kann durch Grenzübergang erledigt werden.

Der Schluß funktioniert auch dann, wenn $u \in M_0$ ist und auf einer Teilmenge E_1 von $E(u>0)$, die als nicht leer vorausgesetzt wird, der Ungleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} \geq \frac{C}{r^2} u \qquad (r \geq r(\varepsilon),\ C = C(\varepsilon) < 0) \tag{7.4.6}$$

genügt.

Diese Teilmenge E_1 soll folgende Eigenschaften besitzen:

1. Ist $P \in E_1$ und liegt P auf der Kugel K_r, $x_1^2 + \cdots + x_n^2 = r^2$, so liegt der Durchschnitt $K_r \cap E$ in E_1.

2. Es sei $\varphi(x)$ die charakteristische Funktion der Menge E_1^* der r-Achse, die alle Radien der K_r enthält, die Punkte von E_1 enthalten.

Dann ist $\int_{E_1^*} \varphi(x) \frac{dx}{x}$ divergent.

Wiederholt man dann die eben durchgeführten Rechnungen, so wird

$$r\left(\frac{m''}{m'} + \frac{1}{r}\right) \geq C_1 \varphi(r)\, m^\varepsilon, \tag{7.4.7}$$

d.h.

$$\frac{d}{d\xi} \log \frac{dm}{d\xi} \geq C_1 \varphi(r)\, m^\varepsilon. \tag{7.4.8}$$

Daraus folgt für jedes $\xi \geq \xi(\varepsilon)$

$$\frac{d^2 m}{d\xi^2} \geq C_2 \varphi \frac{dm^{1+\varepsilon}}{d\xi},$$

und infolgedessen für jedes ξ mit $e^\xi \in E_1^*$

$$\frac{dm}{d\xi} \geq C_1 \varphi(\xi)\, m^{1+\varepsilon}. \tag{7.4.9}$$

Denn sowohl m als auch $dm/d\xi$ sind monoton wachsend und letztere bei geeigneter Wahl von $\xi(\varepsilon)$ positiv.

Aus (6.4.6) erhält man jetzt für jedes $r \in E_1^*$

$$\varphi(r) \frac{dr}{r} \leq C_2 \frac{dm}{m^{1+\varepsilon}}, \tag{7.4.10}$$

woraus dann durch Integration die Behauptung folgt.

Ich will die Zahl der Beispiele der Anwendungsmöglichkeit der Carlemanschen Methode nicht weiter erhöhen. Man könnte natürlich auch manche Probleme aus dem Lindelöfschen Kreis verschärfen, indem man z.B. annimmt, daß die Lindelöfsche Bedingung $u \leq 0$ nicht überall erfüllt ist, sondern auf Gebieten von $x_1 = 0$, in welchen eine unendliche Folge von monoton wachsenden Ringgebieten Platz findet, die sich ins Unendliche erstrecken. Auch könnte man einen Vorstoß zu denjenigen Lindelöfschen Problemen unternehmen, welche durch die allgemeine elliptische partielle Differentialgleichung eröffnet werden. Einen guten Anfang in dieser Richtung hat Huber gemacht (Proc. Amer. Math. Soc. **4**, 852–857, 1953).

Andere Probleme, die sich mit der hier entwickelten Carlemanschen Methode angreifen lassen, bewegen sich in der Richtung des Denjoy-Carleman-Ahlforsschen Satzes. In dieser Richtung liegt, soviel ich weiß, nur ein Ergebnis vor, das ich in einer kurzen Comptes Rendus Note aus dem Jahre 1954 angegeben habe. Erst wenn man diese Resultate hinzunimmt, erhält man einen Einblick über die Menge der Probleme, in die man durch die hier entwickelte Methode vordringen kann. Ich glaube jedoch, daß der Leser durch die hier zur Darstellung gekommenen Auswahl der Anwendungen der verallgemeinerten Carlemanschen Methode einen überzeugenden Eindruck von der Stärke der ursprünglichen Carlemanschen Idee erhalten hat.

Berlin-Dahlem, Mathematisches Institut der Freien Universität.

(Eingegangen am 11. Januar 1955.)

SCHULZ, W.
Math. Zeitschr. Bd. 63, S. 133—144 (1955)

Über Reduzibilität bei gewissen Polynomen und das Tarry-Escottsche Problem.

Zum Gedenken an den 80. Geburtstag meines verehrten Lehrers I. SCHUR am 10. Januar 1955.

Von

WERNER SCHULZ.

1

I. SCHUR befaßte sich wohl als erster mit der Frage, ob Polynome, die an gewissen ganzzahligen Stellen vorgeschriebene Werte annehmen, im Körper der rationalen Zahlen reduzibel oder irreduzibel sind. Die beiden einfachsten Fälle dieser Art, bei denen nämlich die Anzahl der Stellen gleich dem Grad des Polynoms ist und die vorgeschriebenen Werte sämtlich gleich $+1$ bzw. -1 sind, wurden von ihm im Jahre 1908 als Aufgabe gestellt [*1*]. Es gelten hierüber die Sätze:

I. Ein Polynom n-ten Grades mit dem höchsten Koeffizienten 1, das an n ganzzahligen Stellen den Wert 1 annimmt, ist im Körper der rationalen Zahlen irreduzibel, sofern $n \neq 2$ und $\neq 4$ ist. Im Falle $n = 2$ tritt Reduzibilität dann und nur dann ein, wenn die beiden Stellen den Abstand 2 voneinander haben. Im Falle $n = 4$ ergeben sich reduzible Polynome dann und nur dann, wenn die vier Stellen vier aufeinanderfolgende ganze Zahlen sind. In allen Fällen von Reduzibilität ist das zerfallende Polynom das Quadrat eines irreduziblen Polynoms.

II. Ein Polynom n-ten Grades mit dem höchsten Koeffizienten 1, das an n ganzzahligen Stellen den Wert -1 annimmt, ist im Körper der rationalen Zahlen irreduzibel.

Diese Aufgaben sind von SCHUR selbst und später in mehreren Arbeiten von Verfassern, die in Berlin dem SCHURschen Schüler- und Hörerkreis angehörten, nach verschiedenen Seiten hin verallgemeinert worden. Genannt seien die Namen A. BRAUER, R. BRAUER, H. HOPF, H. ILLE, U. WEGNER. Nähere Angaben hierüber findet man in der Arbeit, die die Herren ALFRED und RICHARD BRAUER in dieser Zeitschrift vor 20 Jahren SCHUR zum 60. Geburtstag widmeten [*2*].

Auch außerhalb des SCHURschen Kreises fand der hier angeschnittene Fragenkomplex Bearbeiter. Einschlägige Untersuchungen von J. WESTLUND, G. PÓLYA, H. L. DORWART, O. ORE, L. WEISNER sind in [*2*] zitiert. Besonderes Interesse hat hierbei ein Fall gefunden, den zuerst PÓLYA [*3*] aus einem allgemeineren Kriterium erhielt:

III. Es sei $n \geq 17$. Wenn unter den Werten des ganzzahligen Polynoms $f(x)$ vom Grade n dieselbe Primzahl, positiv oder negativ genommen, n-mal vorkommt, dann ist $f(x)$ entweder irreduzibel oder das Produkt zweier irreduzibler Faktoren gleichen Grades.

Daß die Zahl 17, die in diesem Satz auftritt, durch die kleinere 7 ersetzt werden kann, zeigte A. BRAUER [4]. Für $n=6$ dagegen ist der PÓLYAsche Satz nicht mehr richtig, wie ein in [4] angeführtes Beispiel zeigt: Das Polynom 6. Grades mit dem höchsten Koeffizienten 1, das an 6 aufeinanderfolgenden ganzzahligen Stellen den Wert -5 annimmt, ist durch ein quadratisches Polynom mit ganzen rationalen Koeffizienten teilbar.

Unabhängig von A. BRAUER und fast zur gleichen Zeit wie dieser befaßten sich auch H. L. DORWART und O. ORE [5] mit dem Satz III.

2

Für den Satz III bleibt die Frage zu entscheiden, ob im Falle gerader Gradzahlen $n=2m \geq 8$ von $f(x)$ der Zerfall in zwei Faktoren m-ten Grades eintreten kann, und mit dieser Frage sollen sich die folgenden Betrachtungen weiter beschäftigen.

Für $n=8$, also $m=4$, gab schon PÓLYA in [3] ein Beispiel für Reduzibilität an:

IV. Das Polynom 8. Grades mit dem höchsten Koeffizienten 1, das an den 8 Stellen ± 3, ± 5, ± 7, ± 11 den Primzahlwert 2879 annimmt, ist reduzibel.

Für den Fall $n=10$ fand ich ein reduzibles Polynom, das A. BRAUER in [4] anführt und das in einer gegenüber dort etwas geänderten Formulierung lautet:

V. Das Polynom 10. Grades mit dem höchsten Koeffizienten 1, das an den 10 Stellen ± 2, ± 4, ± 7, ± 9, ± 10 den Primzahlwert 10079 annimmt, ist reduzibel.

In Abschnitt 4 wird sich zeigen, daß bis zum Grad $2m=20$ die Existenz reduzibler Polynome der betrachteten Art gesichert ist. Darüber hinaus ist die Frage nach der Möglichkeit eines Zerfalls bis heute unerledigt.

Wichtig sind folgende beiden Hilfssätze, die sich aus den Betrachtungen in [4] und [5] ergeben:

VI. Ein ganzzahliges Polynom, das an mindestens fünf verschiedenen ganzzahligen Stellen den Wert ± 1 annimmt, nimmt an diesen Stellen entweder überall den Wert $+1$ oder überall den Wert -1 an.

VII. Ein ganzzahliges Polynom, das an mindestens sechs verschiedenen ganzzahligen Stellen den Wert $\pm p$ annimmt (p = Primzahl), nimmt an diesen Stellen entweder überall den Wert $+p$ oder überall den Wert $-p$ an.

Diese beiden Hilfssätze bilden die Grundlage für eine genauere Einsicht in die Art des Zerfalls bei geraden Gradzahlen. Folgende Erkenntnis wird hierüber in [4] gewonnen:

VIII. Für ein ganzzahliges Polynom vom Grade $2m \geq 8$, das an $2m$ ganzzahligen Stellen den Wert $\pm p$ annimmt, kann ein Zerfall in zwei Faktoren m-ten Grades nur in der Weise eintreten, daß die Summe oder die Differenz der beiden Faktoren konstant ist.

3

Es soll nun über den Satz VIII hinausgehend eine genauere Aussage hinsichtlich des Zerfalls in zwei Faktoren bewiesen werden, die in etwas anderer Form bereits in meiner Dissertation [6] enthalten ist. Die Polynome vom Grade $2m \geq 8$, die auf ihre Reduzibilität hin zu untersuchen sind, können auf Grund von VII in der Form

$$f(x) = A(x - c_1)(x - c_2) \dots (x - c_{2m}) \pm p \quad (A > 0) \tag{1}$$

angesetzt werden. Existiert eine Zerlegung

$$f(x) = g(x)\, h(x), \tag{2}$$

$$\operatorname{Grad} g(x) = \operatorname{Grad} h(x) = m, \tag{3}$$

so darf angenommen werden, daß für $g(x)$ und $h(x)$ die Koeffizienten von x^m positiv sind. Von den beiden in VIII genannten Fällen muß dann der eintreten, daß die Differenz $g(x) - h(x)$ konstant ist. Zugleich zeigt sich, da die höchsten Koeffizienten von $g(x)$ und $h(x)$ übereinstimmen und ihr Produkt gleich A ist, daß A eine Quadratzahl sein muß.

Von den $2m$ (nach Voraussetzung voneinander verschiedenen) Stellen c_ν kann angenommen werden, daß sie nach wachsender Größe geordnet sind. Ferner kann man sie mittels einer linearen Transformation von x, die an den Reduzibilitätsverhältnissen von $f(x)$ nichts ändert, so legen, daß die erste Stelle auf Null fällt. Dann gelten also die Bedingungen

$$0 = c_1 < c_2 < \dots < c_{2m}. \tag{4}$$

Es wird behauptet:

IX. *Wenn das Polynom* (1) *vom Grade* $2m \geq 8$ *reduzibel ist und für die Stellen* c_ν *die Bedingungen* (4) *gelten, dann haben die beiden Faktoren* $g(x)$ *und* $h(x)$ *von* $f(x)$ *die Form*[1])

$$\begin{cases} g(x) = A'(x - a_1)(x - a_2) \dots (x - a_m) \pm (-1)^m \\ \quad\; = A'(x - b_1)(x - b_2) \dots (x - b_m) - (-1)^m p, \end{cases} \tag{5}$$

$$\begin{cases} h(x) = A'(x - a_1)(x - a_2) \dots (x - a_m) + (-1)^m p \\ \quad\; = A'(x - b_1)(x - b_2) \dots (x - b_m) \mp (-1)^m. \end{cases} \tag{6}$$

Hierbei ist

$$A' = \sqrt{A}, \tag{7}$$

[1]) In Ausdrücken, die doppelte Vorzeichen besitzen, gilt das obere oder untere Vorzeichen, je nachdem ob $f(x)$ an den Stellen c_ν den Wert $+p$ oder $-p$ annimmt.

die Werte a_μ und b_μ sind in ihrer Gesamtheit die Werte c_ν, und zwar in folgender Anordnung:

$$\begin{cases} a_{2\varkappa-1} = c_{4\varkappa-3}, & a_{2\varkappa} = c_{4\varkappa}, \\ b_{2\varkappa-1} = c_{4\varkappa-2}, & b_{2\varkappa} = c_{4\varkappa-1}, \end{cases} \tag{8}$$

und es gilt

$$p = A' b_1 b_2 \dots b_m \mp 1 = A' \prod_{\varkappa} c_{4\varkappa-2} c_{4\varkappa-1} \mp 1. \tag{9}$$

Beweis. Die Richtigkeit von (7) wurde bereits oben erkannt.
Aus (1) und (2) ergibt sich

$$g(c_\nu)\, h(c_\nu) = \pm p. \tag{10}$$

Alle $g(c_\nu)$ und $h(c_\nu)$ können daher nur die vier Werte $+1$, -1, $+p$, $-p$ besitzen, und die Stellen c_ν können hiernach in vier Sorten eingeteilt werden, derart daß

$$k \text{ Stellen mit } g(c_\nu) = +1, \quad h(c_\nu) = \pm p, \tag{11}$$

$$l \text{ Stellen mit } g(c_\nu) = -1, \quad h(c_\nu) = \mp p, \tag{12}$$

$$r \text{ Stellen mit } g(c_\nu) = +p, \quad h(c_\nu) = \pm 1, \tag{13}$$

$$s \text{ Stellen mit } g(c_\nu) = -p, \quad h(c_\nu) = \mp 1 \tag{14}$$

existieren. Mit derselben Überlegung wie beim Beweis des Satzes VIII in [4], wobei für $2m \geqq 10$ nur der Satz VI gebraucht wird und für $2m = 8$ eine speziellere Betrachtung zum Ziele führt, erkennt man, daß nur folgende vier Fälle möglich sind:

(a) $$k = r = m, \quad l = s = 0,$$

(b) $$k = s = m, \quad l = r = 0,$$

(c) $$l = r = m, \quad k = s = 0,$$

(d) $$l = s = m, \quad k = r = 0.$$

Aus (11) bis (14) ergibt sich, daß folgende Polynome in den Fällen (a) bis (d), deren Grad höchstens gleich m ist, an $2m$ Stellen denselben Wert annehmen und daher konstant sein müssen:

(a) $$g(x) \pm h(x) = \quad p + 1,$$

(b) $$g(x) \mp h(x) = -p + 1,$$

(c) $$g(x) \mp h(x) = \quad p - 1,$$

(d) $$g(x) \pm h(x) = -p - 1.$$

Da jedoch nach den zu Beginn dieses Abschnitts gemachten Voraussetzungen die Summe $g(x) + h(x)$ nicht konstant sein kann, so zeigt sich, daß die Fälle (a) und (d) nicht eintreten können, wenn $f(c_\nu) = +p$, und die Fälle (b) und (c) nicht, wenn $f(c_\nu) = -p$ ist. Für $g(x)$ und $h(x)$ bleiben hiernach sowohl im Falle

$f(c_\nu) = +p$ als auch im Falle $f(c_\nu) = -p$ nur zwei Möglichkeiten, die, wie man sich leicht überzeugt, in der Form

$$(15)\qquad \begin{cases} g(x) = A'(x-a_1)(x-a_2)\dots(x-a_m) + \varepsilon \\ \quad = A'(x-b_1)(x-b_2)\dots(x-b_m) \mp \varepsilon p, \end{cases}$$

$$(16)\qquad \begin{cases} h(x) = A'(x-a_1)(x-a_2)\dots(x-a_m) \pm \varepsilon p \\ \quad = A'(x-b_1)(x-b_2)\dots(x-b_m) - \varepsilon \end{cases}$$

zusammengefaßt werden können, wobei $\varepsilon = +1$ oder -1 sein kann.

Unter den a_μ und b_μ ist nach (4) ein Wert gleich Null. Man kann ohne Beschränkung der Allgemeinheit annehmen, daß bei $g(x)$ und $h(x)$ die Benennungen so gewählt sind, daß $a_1 = c_1 = 0$ ist und die a_μ und b_μ nach wachsender Größe geordnet sind:

$$(17)\qquad 0 = a_1 < a_2 < \dots < a_m,$$

$$(18)\qquad 0 < b_1 < b_2 < \dots < b_m.$$

Für das Polynom (15) folgt an der Stelle $x = 0$ aus (17)

$$(19)\qquad \varepsilon = (-1)^m A' b_1 b_2 \dots b_m \mp \varepsilon p,$$

d.h.

$$(20)\qquad p = \pm (-1)^m \varepsilon A' b_1 b_2 \dots b_m \mp 1.$$

Da aber p positiv ist, erhält man aus (20) und (18)

$$(21)\qquad \varepsilon = \pm (-1)^m$$

und demnach

$$(22)\qquad p = A' b_1 b_2 \dots b_m \mp 1.$$

Wegen (15), (16) und (21) haben also die Faktoren von $f(x)$ die Form (5) und (6).

Als nächstes betrachte man die Ableitung des Polynoms $g(x)$, die ein Polynom $(m-1)$-ten Grades ist. Die Nullstellen seien $\xi_1, \xi_2, \dots, \xi_{m-1}$. Da $g(x)$ an allen Stellen a_μ denselben Wert annimmt, muß zwischen je zwei aufeinanderfolgenden a ein ξ liegen. Dasselbe gilt für je zwei aufeinanderfolgende b. Daher erhält man wegen (17) und (18)

$$(23)\qquad 0 = a_1 < \xi_1 < a_2 < \xi_2 < \dots < \xi_{m-1} < a_m,$$

$$(24)\qquad 0 < b_1 < \xi_1 < b_2 < \xi_2 < \dots < \xi_{m-1} < b_m.$$

Ferner ist zu bemerken, daß

$$(25)\qquad \begin{cases} g(x) \neq g(a_1) = g(a_2) = \dots = g(a_m) & \text{für} \quad x \neq a_\mu, \\ g(x) \neq g(b_1) = g(b_2) = \dots = g(b_m) & \text{für} \quad x \neq b_\mu \end{cases}$$

ist, da $g(x)$ vom Grade m ist.

Für die Lage von ξ_1 ergibt sich aus (23) und (24) notwendig

$$a_1 < b_1 < \xi_1.$$

Die Werte, die $g(x)$ im Intervall $a_1 \leq x \leq \xi_1$ annimmt, muß es aus Stetigkeitsgründen auch im Intervall $\xi_1 \leq x \leq a_2$ annehmen, und daher muß wegen (25)

$$\xi_1 < b_2 < a_2 < \xi_2$$

sein. Dieselbe Schlußweise, angewandt auf die Intervalle $b_2 \leq x \leq \xi_2$ und $\xi_2 \leq x \leq b_3$, liefert

$$\xi_2 < a_3 < b_3 < \xi_3.$$

Setzt man diese Betrachtung fort, so bekommt man als Verteilung der Stellen a_μ und b_μ:

$$a_1 < b_1 < b_2 < a_2 < a_3 < b_3 < b_4 < a_4 < \cdots. \tag{26}$$

Der Vergleich von (26) mit (4) ergibt daher die Beziehungen (8).

Aus (22) und (8) erkennt man schließlich die Richtigkeit von (9).

Damit ist Satz IX vollständig bewiesen.

4

Die Tatsache, daß die Differenz der Polynome (5) und (6) eine Konstante ist, hat zur Folge, daß auch die beiden Polynome

$$G(x) = (x - a_1)(x - a_2) \dots (x - a_m) \tag{27}$$

und

$$H(x) = (x - b_1)(x - b_2) \dots (x - b_m) \tag{28}$$

sich nur um eine Konstante unterscheiden. Unter Voraussetzung der in (17) angenommenen Normierung $a_1 = 0$ folgt aus (5) und (9)

$$H(x) - (-1)^m b_1 b_2 \dots b_m = G(x). \tag{29}$$

Hat man umgekehrt zwei Polynome $G(x)$ und $H(x)$ von der Form (27) und (28) mit lauter voneinander verschiedenen ganzzahligen Nullstellen, derart daß $G(x) - H(x)$ konstant ist, so gehören dazu unendlich viele Polynome $f(x)$ der Form (1), die reduzibel sind. Denn wenn es Polynome $G(x)$ und $H(x)$ der verlangten Art gibt, so kann man durch eine lineare Transformation von x und geeignete Wahl der Bezeichnungen erreichen, daß die Stellen a_μ und b_μ die Bedingungen (17) und (18) erfüllen. Dies sei also von vornherein für (27) und (28) angenommen. Wählt man dann A' so, daß $A' b_1 b_2 \dots b_m \mp 1$ eine Primzahl p wird, was nach dem Satz über die Primzahlen in einer arithmetischen Progression auf unendlich viele Arten möglich ist, so haben die Polynome

$$f(x) = A'^2 G(x) H(x) \pm p \tag{30}$$

die gewünschte Eigenschaft. Die Frage nach der Existenz von reduziblen Polynomen (1) ist also gleichbedeutend mit der Frage nach der Existenz von

Polynomen (27) und (28), die nur in ihren absoluten Gliedern voneinander verschieden sind, d.h. mit der Frage nach der Existenz von Systemen ganzer rationaler Zahlen $a_1, a_2, \ldots, a_m$ und $b_1, b_2, \ldots, b_m$, deren erste $m-1$ elementarsymmetrischen Funktionen übereinstimmen. Auf Grund der NEWTONschen Relationen, die zwischen den elementarsymmetrischen Funktionen und den Potenzsummen von Größen bestehen, ist diese Frage aber gleichwertig mit der Aufgabe, Systeme ganzer rationaler Zahlen a_μ und b_μ zu finden, deren erste $m-1$ Potenzsummen übereinstimmen:

$$(31) \qquad a_1^k + a_2^k + \cdots + a_m^k = b_1^k + b_2^k + \cdots + b_m^k \quad (k = 1, 2, \ldots, m-1).$$

Für das Gleichungssystem (31) wird auch kurz die Schreibweise

$$(32) \qquad a_1, a_2, \ldots, a_m \overset{m-1}{=} b_1, b_2, \ldots, b_m$$

verwendet.

Die Aufgabe, ganzzahlige Lösungen des Gleichungssystems

$$(33) \qquad a_1, a_2, \ldots, a_m \overset{l}{=} b_1, b_2, \ldots, b_m$$

zu finden, wobei l eine Zahl unterhalb m ist, hat nun, worauf DORWART [7] aufmerksam machte, unabhängig von ihrem Zusammenhang mit der Frage nach der Reduzibilität der Polynome (1) seit langem zahlreiche Bearbeiter gefunden, die sich mit ihr teils wegen ihrer Kuriosität, teils wegen ihres Zusammenhanges mit anderen Aufgaben, wie z.B. der Aufstellung schnell konvergierender Reihen zur numerischen Berechnung von Logarithmen beschäftigten. Für dieses Problem hat sich nach einem Vorschlag von L. E. DICKSON der Name „TARRY-ESCOTTsches Problem" eingebürgert. In dem Fall $l = m-1$, der für die hier angestellten Betrachtungen von Interesse ist, pflegt man von dem „idealen TARRY-ESCOTTschen Problem" zu sprechen. Literaturangaben hierüber sind an den unter [8] zitierten Stellen zu finden.

Der Engländer ESCOTT [9] und der Franzose TARRY [10a] gaben 1910 bzw. 1912 ein Verfahren an, um aus zwei Systemen von Zahlen, deren erste l elementarsymmetrischen Funktionen oder Potenzsummen übereinstimmen, zwei Systeme von Zahlen zu gewinnen, deren erste $l+1$ elementarsymmetrischen Funktionen oder Potenzsummen übereinstimmen. Die Formulierung bei ESCOTT, die in ähnlicher Gestalt allerdings schon fast 100 Jahre vorher von KNIGHT [11] ausgesprochen wurde, lautet:

X. Wenn für zwei Polynome $f_1(x)$ und $f_2(x)$ die ersten r Koeffizienten, beginnend mit dem höchsten, übereinstimmen, so stimmen für die Polynome $f_1(x) f_2(x+d)$ und $f_2(x) f_1(x+d)$ bei beliebigem d die ersten $r+1$ Koeffizienten überein.

TARRY spricht den Satz so aus:

XI. Wenn für $a_1, \ldots, a_s$ und $b_1, \ldots, b_s$ die ersten l Potenzsummen übereinstimmen, so stimmen für $a_1, \ldots, a_s, b_1+d, \ldots, b_s+d$ und $b_1, \ldots, b_s, a_1+d, \ldots, a_s+d$ die ersten $l+1$ Potenzsummen überein.

Der Satz XI kann dazu benutzt werden, um fortlaufend Zahlensysteme zu konstruieren, bei denen immer mehr Potenzsummen übereinstimmen.

Hierbei wird es sich jedoch im allgemeinen nur um Lösungen des allgemeinen, nicht des idealen TARRY-ESCOTTschen Problems handeln. Nur unter besonders günstigen Umständen, nämlich dann, wenn genügend viele Differenzen $a_\varkappa - a_\lambda$ und $b_\varkappa - b_\lambda$ den Wert d besitzen, wird man zu einer idealen Lösung gelangen. So ergibt sich z.B. für $m=7$ eine Lösung von (32), wenn man von

$$\begin{matrix} 0,\ 7,\ 11,\ 17,\ 18,\ 24,\ 28,\ 35 \\ \overset{1}{=}\ 1,\ 8,\ 12,\ 15,\ 20,\ 23,\ 27,\ 34 \end{matrix}$$

ausgeht und den TARRYschen Satz der Reihe nach mit $d=7,\ 11,\ 13,\ 17,\ 19$ anwendet. Das führt auf

$$\begin{matrix} 0,\ 18,\ 27,\ 58,\ 64,\ 89,\ 101 \\ \overset{6}{=}\ 1,\ 13,\ 38,\ 44,\ 75,\ 84,\ 102. \end{matrix}$$

Dieses Beispiel findet sich schon bei ESCOTT ([9], S. 152; vgl. auch [8e] S. 331).

Für $m=8$ erhält man eine Lösung, wenn man mit

$$0,\ 8 \overset{1}{=} 3,\ 5$$

beginnt und Satz XI mit $d=1,\ 2,\ 7,\ 8,\ 11,\ 13$ anwendet. Dann entsteht

$$\begin{matrix} 0,\ 4,\ \ 9,\ 23,\ 27,\ 41,\ 46,\ 50 \\ \overset{7}{=}\ 1,\ 2,\ 11,\ 20,\ 30,\ 39,\ 48,\ 49. \end{matrix}$$

Dieses Beispiel ist ebenfalls schon alt; TARRY fand es als erster ([10b]; vgl. auch [8b], S. 616).

Später wurden noch für $m=9$ und $m=10$ Lösungen von (32) gefunden. A. LETAC gab 1942 zwei Zahlenbeispiele für den Fall $m=9$ an und zeigte einen Weg, der zu unendlich vielen Lösungen für den Fall $m=10$ führt (vgl. [8f], S. 48, 54, 56). Für $m>10$ sind bisher keine Lösungen von (32) in ganzen rationalen Zahlen bekannt.

5

Aus (27), (28) und (29) erhält man sofort weitere Bedingungen, die die Zahlen a_μ und b_μ notwendig erfüllen müssen, wenn sie eine Lösung des idealen TARRY-ESCOTTschen Problems ergeben sollen. Ist q eine Primzahl, die in einer der Zahlen $b_1, b_2, \ldots, b_m$ aufgeht, so werden wegen (29) die Polynome $G(x)$ und $H(x)$ mod q einander identisch. Daher müssen die Nullstellen von (27) und (28) übereinstimmen. Es gilt also folgender Satz, der in etwas umständlicherer Form schon in [6], S. 148 bewiesen ist:

XII. Ist q Primteiler einer der Zahlen b_μ, so stimmen die Zahlen $a_1, a_2, \ldots, a_m$ und $b_1, b_2, \ldots, b_m$ mod q überein.

Folgerungen aus diesem Satz, auf die hier nicht eingegangen werden soll, finden sich in [6].

Man kann auch, wie in [6] gezeigt ist, Primzahlen angeben, die als Teiler der b_μ vorkommen müssen. Zum Beispiel ist sicher jede in m nicht enthaltene

Primzahl q unterhalb m Teiler eines b_μ und muß in dem Produkt $b_1 b_2 \dots b_m$ mindestens in der $\left[\frac{m+q-1}{q}\right]$-ten Potenz aufgehen. Von Nutzen ist beim Beweis dieses Satzes die Betrachtung der m-reihigen quadratischen Matrix

$$(a_\varkappa - a_\lambda) \quad (\varkappa, \lambda = 1, 2, \dots, m). \tag{34}$$

Da aus (27), (28) und (29) für jedes $a_\varkappa$ und jedes b_λ $(\varkappa, \lambda = 1, 2, \dots, m)$

$$\begin{cases} H(a_\varkappa) = (a_\varkappa - b_1)(a_\varkappa - b_2) \dots (a_\varkappa - b_m) = (-1)^m b_1 b_2 \dots b_m, \\ G(b_\lambda) = (b_\lambda - a_1)(b_\lambda - a_2) \dots (b_\lambda - a_m) = -(-1)^m b_1 b_2 \dots b_m \end{cases} \tag{35}$$

folgt, so wird in der Matrix (34) das Produkt aus den Elementen jeder Zeile gleich $(-1)^m b_1 b_2 \dots b_m$ und das Produkt aus den Elementen jeder Spalte gleich $-b_1 b_2 \dots b_m$ (vgl. [6], S. 151 und [8d]). Das ideale TARRY-ESCOTTsche Problem kann daher auch so formuliert werden:

Es ist eine m-reihige quadratische Matrix $(n_{\varkappa\lambda})$ mit folgenden Eigenschaften gesucht:

1. Die Zahlen der ersten Zeile sind negative ganze rationale Zahlen

$$n_{1\lambda} = -b_\lambda < 0 \quad (\lambda = 1, 2, \dots, m).$$

2. Die übrigen Zeilen entstehen aus der ersten Zeile durch Addition positiver Zahlen $a_2, a_3, \dots, a_m$:

$$n_{\varkappa\lambda} = a_\varkappa - b_\lambda \quad (\varkappa = 2, 3, \dots, m).$$

3. Die Produkte $n_{\varkappa 1} n_{\varkappa 2} \dots n_{\varkappa m}$ haben für jedes $\varkappa = 1, 2, \dots, m$ denselben Wert $(-1)^m N$.

4. Die Produkte $n_{1\lambda} n_{2\lambda} \dots n_{m\lambda}$ haben für jedes $\lambda = 1, 2, \dots, m$ denselben Wert $-N$.

6

Da es bisher nicht gelungen ist zu entscheiden, ob das ideale TARRY-ESCOTTsche Problem (32) für beliebiges m in ganzen rationalen Zahlen lösbar ist, und da die Ausführungen in Abschnitt 5 zeigen, daß Zahlen, die eine solche Lösung liefern, eine ganze Reihe von Bedingungen notwendig erfüllen müssen, scheint es nicht ohne Interesse, die Frage zu stellen, ob das Gleichungssystem (32) in anderen als ganzen rationalen Zahlen lösbar ist.

Trivial ist der Fall, daß für die a_μ und b_μ *beliebige reelle oder komplexe Zahlen* zugelassen werden. Dann ist (32) für jedes m lösbar. Man braucht nur zwei Polynome m-ten Grades zu nehmen, deren Differenz konstant ist. Die Nullstellen dieser beiden Polynome ergeben eine Lösung von (32).

Auch daß (32) für jedes m Lösungen in *reellen* Zahlen zuläßt, erkennt man sofort. Es seien nämlich $a_1, a_2, \dots, a_m$ reelle Zahlen, die (17) erfüllen. Mit diesen Größen bilde man nach (27) das Polynom $G(x)$. Die Nullstellen der Ableitung von $G(x)$ sind voneinander verschiedene reelle Zahlen $\xi_1, \xi_2, \dots, \xi_{m-1}$, und es ist, wenn die ξ_ν nach wachsender Größe geordnet sind,

$$\operatorname{sgn} G(\xi_{2\varkappa-1}) = (-1)^{m-1} \quad \left(\varkappa = 1, 2, \dots, \left[\frac{m}{2}\right]\right).$$

Setzt man

$$\gamma = \operatorname*{Min}_{x} |G(\xi_{2x-1})|,$$

so sind die Nullstellen jedes Polynoms $G(x) + (-1)^m C$ mit $0 < C < \gamma$ voneinander verschiedene positive reelle Zahlen $b_1, b_2, \ldots, b_m$, und die Zahlen a_μ und b_μ sind eine Lösung von (32).

Man sieht übrigens wie beim Beweis des Satzes IX, daß der Größe nach geordnete reelle Zahlen a_μ und b_μ, die Lösung von (32) sind, die Ungleichungen (26) erfüllen müssen.

Als letztes soll die Frage betrachtet werden, ob das *Kongruenzensystem*

$$a_1, a_2, \ldots, a_m \overset{m-1}{\equiv} b_1, b_2, \ldots, b_m \pmod{M} \tag{36}$$

nichttriviale Lösungen besitzen kann, d.h. solche, bei denen die b_μ nicht bloß eine Permutation der $a_\mu \pmod M$ darstellen. Nach Satz XII erhält man jedenfalls aus einer Lösung des idealen TARRY-ESCOTTschen Problems in ganzen rationalen Zahlen a_μ, b_μ keine nichttriviale Lösung des zugehörigen Problems mod M, wenn M Primteiler einer der Zahlen b_μ ist.

Beispiele, die zeigen, daß (36) für ungerade Primzahlmoduln $M = p$ Lösungen besitzen kann, sind folgende:

1. $m = p - 1, \quad a_\mu \equiv \mu, \quad b_\mu \equiv 0 \pmod p, \quad \mu = 1, 2, \ldots, p-1,$

also:

$$1, 2, \ldots, p-1 \overset{p-2}{\equiv} 0 \pmod p. \tag{37}$$

Die Summe der Elemente der linken Seite von (37) ist nämlich durch p teilbar, und die zweite, dritte, ..., $(p-2)$-te Potenz der Elemente ergibt stets nur eine Permutation dieser Elemente. Man kann auch so schließen: Die Polynome

$$(x-1)(x-2)\ldots(x-p+1) \equiv x^{p-1} - 1 \pmod p \quad \text{und} \quad x^{p-1}$$

unterscheiden sich mod p nur um eine Konstante.

2. $m = \frac{p-1}{2}, \quad a_\mu \equiv r_\mu, \quad b_\mu \equiv 0 \pmod p, \quad \mu = 1, 2, \ldots, \frac{p-1}{2},$

wobei die r_μ ein vollständiges System quadratischer Reste mod p bilden. Es ist

$$r_1, r_2, \ldots, r_{\frac{p-1}{2}} \overset{\frac{p-3}{2}}{\equiv} 0 \pmod p, \tag{38}$$

da die Polynome

$$\prod_\mu (x - r_\mu) \equiv x^{\frac{p-1}{2}} - 1 \pmod p \quad \text{und} \quad x^{\frac{p-1}{2}}$$

sich mod p nur um eine Konstante unterscheiden.

3. $m = \frac{p-1}{2}, \quad a_\mu \equiv n_\mu, \quad b_\mu \equiv 0 \pmod p, \quad \mu = 1, 2, \ldots, \frac{p-1}{2},$

wobei die n_μ ein vollständiges System quadratischer Nichtreste mod p bilden. Es ist

$$n_1, n_2, \ldots, n_{\frac{p-1}{2}}^{\frac{p-3}{2}} \equiv 0 \pmod{p}, \tag{39}$$

wie die Betrachtung der Polynome

$$\prod_\mu (x - n_\mu) \equiv x^{\frac{p-1}{2}} + 1 \pmod{p} \quad \text{und} \quad x^{\frac{p-1}{2}}$$

zeigt.

4. Aus (38) und (39) folgt

$$r_1, r_2, \ldots, r_{\frac{p-1}{2}}^{\frac{p-3}{2}} \equiv n_1, n_2, \ldots, n_{\frac{p-1}{2}} \pmod{p}, \tag{40}$$

d.h. die Zahlen $1, 2, \ldots, p-1$ lassen sich so in zwei Mengen von je $m = \frac{p-1}{2}$ Zahlen zerlegen, daß eine Lösung des idealen TARRY-ESCOTTschen Kongruenzensystems (36) entsteht.

5. Für zusammengesetzte Moduln M sind entsprechende Schlüsse nicht zulässig. Man habe z.B. zwei Polynome $G(x)$ und $H(x)$ vom Grade m mit dem höchsten Koeffizienten 1, deren Differenz mod M eine Konstante ist, und man kenne für $G(x)$ und $H(x)$ je m Nullstellen $a_1, a_2, \ldots, a_m$ und $b_1, b_2, \ldots, b_m$ mod M. Dann müssen nicht notwendig die Zahlen a_μ und b_μ Lösung eines idealen TARRY-ESCOTTschen Kongruenzensystems (36) sein. Denn die Produkte $\prod_\mu (x - a_\mu)$ und $\prod_\mu (x - b_\mu)$ brauchen nicht notwendig mod M mit $G(x)$ und $H(x)$ übereinzustimmen. Ein Beispiel stellen etwa die Polynome

$$G(x) = x^{\varphi(M)} - 1, \quad H(x) = x^{\varphi(M)}, \quad \varphi(M) = \text{EULERsche Funktion},$$

dar. Die zu M teilerfremden Zahlen $a_1, a_2, \ldots, a_{\varphi(M)}$ eines vollständigen Restsystems mod M sind die Nullstellen von $G(x)$. Die Zahlen a_μ und $b_1 = b_2 = \cdots = b_{\varphi(M)} = 0$ ergeben im allgemeinen aber keine Lösung von (36) mit $m = \varphi(M)$.

Man darf auch nicht umgekehrt aus der Tatsache, daß für zwei Zahlensysteme $a_1, a_2, \ldots, a_m$ und $b_1, b_2, \ldots, b_m$ die ersten, zweiten, ..., $(m-1)$-ten Potenzsummen mod M übereinstimmen, den Schluß ziehen, daß sich die Polynome $\prod_\mu (x - a_\mu)$ und $\prod_\mu (x - b_\mu)$ nur um eine Konstante mod M unterscheiden. Denn die NEWTONschen Relationen liefern für einen zusammengesetzten Modul aus dem Übereinstimmen der Potenzsummen der a_μ und b_μ nicht notwendig das Übereinstimmen der entsprechenden elementarsymmetrischen Funktionen. Als Beispiel sei angeführt, daß für

$$a_\mu \equiv 1, \quad b_\mu \equiv 0 \pmod{M} \quad (\mu = 1, 2, \ldots, M) \tag{41}$$

alle Potenzsummen mod M einander kongruent werden, daß die Zahlen (41) also sicher eine Lösung von (36) mit $m = M$ ergeben. Die Polynome $(x-1)^M$ und x^M unterscheiden sich jedoch mod M, wenn M keine Primzahl ist, nicht bloß um eine Konstante.

Literatur.

[1] (a) SCHUR, I.: Aufgabe 226. Arch. Math. u. Phys. (3) **13**, 367 (1908). — (b) FLÜGEL, W.: Lösung der Aufgabe 226. Arch. Math. u. Phys. (3) **15**, 271—272 (1909). — (c) PÓLYA, G., u. G. SZEGÖ: Aufgaben und Lehrsätze aus der Analysis II. Berlin 1925. VIII. Abschnitt, Aufg. 121 und 122. — [2] BRAUER, A., u. R. BRAUER: Über Irreduzibilitätskriterien von I. SCHUR und G. PÓLYA. Math. Z **40**, 242—265 (1935). — [3] PÓLYA, G.: Verschiedene Bemerkungen zur Zahlentheorie. Jahresber. Dtsch. Math.-Vereinigung **28**, 31—40, insbesondere 34—36 (1919). — [4] BRAUER, A.: Bemerkungen zu einem Satze von Herrn G. PÓLYA. Jahresber. Dtsch. Math.-Vereinigung **43**, 124—129 (1934). — [5] DORWART, H. L., and O. ORE: Criteria for the irreducibility of polynomials. Ann. of Math. **34**, 81—94 (1933). — [6] SCHULZ, W.: Reduzibilität, Irreduzibilität und Affektfreiheit bei gewissen Klassen von Polynomen. Schriften des Math. Sem. und des Inst. f. angew. Math. Universität Berlin **3**, 117—154, insbesondere 144—154 (1937). — [7] DORWART, H. L.: Concerning certain reducible polynomials. Duke Math. J. **1**, 70—73 (1935). — [8] (a) DICKSON, L. E.: History of the Theory of Numbers, Bd. II, S. 705—713. Washington 1920. — (b) DORWART, H. L., and O. E. BROWN: The Tarry-Escott problem. Amer. Math. Monthly **44**, 613—626 (1937). — (c) CHERNICK, J.: Ideal solutions of the Tarry-Escott problem. Amer. Math. Monthly **44**, 626—633 (1937). — (d) BURCHNALL, J. L., and T. W. CHAUNDY: A type of „magic square" in Tarry's problem. Quart. J. of Math. (Oxford series) **8**, 119—130 (1937). — (e) HARDY, G. H., and E. M. WRIGHT: An Introduction to the Theory of Numbers, S. 328—331. Oxford 1938. — (f) GLODEN, A.: Mehrgradige Gleichungen. Groningen 1944. — (g) DORWART, H. L.: Sequences of ideal solutions in the Tarry-Escott problem. Bull. Amer. Math. Soc. **53**, 381—391 (1947). — (h) GLODEN, A., et G. PALAMÀ: Bibliographie des multigrades avec quelques notices biographiques. Luxembourg 1948. — [9] ESCOTT, E. B.: Logarithmic series. Quart. J. of Math. **41**, 141—156 (1910). — [10] (a) TARRY, G.: Interméd. des Math. **19**, 219—222 (1912). — (b) Interméd. des Math. **20**, 68—70 (1913). — [11] KNIGHT, T.: On the construction of logarithmic tables. Phil. Trans. Royal Soc. London **117**, 217—233 (1817). Vgl. hierzu [8h], S. 21.

Braunschweig, Gneisenaustraße 1.

(Eingegangen am 12. Januar 1955.)

FRUCHT, R.
Math. Zeitschr. Bd. 63, S. 145—155 (1955)

Zur Darstellung endlicher Abelscher Gruppen durch Kollineationen.

(Zum 80. Geburtstag von I. SCHUR.)

Von

ROBERT FRUCHT.

Einleitung.

In zwei klassischen Arbeiten [3], [4][1]) hat I. SCHUR eine Theorie der Darstellung endlicher Gruppen durch *gebrochene* lineare Substitutionen — oder wie man unter Verwendung eines geometrischen Ausdruckes kürzer sagen kann, durch Kollineationen — entwickelt. Auf eine Zusammenfassung dieser Theorie können wir hier verzichten, da SCHUR selbst eine solche Zusammenfassung in einer späteren Arbeit ([5], § 1) gegeben hat, in der er seine Theorie auf die symmetrische und die alternierende Gruppe angewendet hat. Außerdem ist für das Verständnis der vorliegenden Arbeit eine genauere Kenntnis der SCHURschen Darstellungstheorie nicht erforderlich[2]), da von ihr kein Gebrauch gemacht werden soll. In der Tat wäre es noch bei Lebzeiten SCHURs Wunsch gewesen, daß einer seiner Schüler eine Vereinfachung der Theorie unternommen hätte; mir ist aber nur eine Arbeit der Herren ASANO und SHODA [1] bekannt, in der zwei Sätze aus der SCHURschen Theorie etwas einfacher bewiesen werden. Da es also scheint, daß eine wesentliche Vereinfachung der SCHURschen Darstellungstheorie bisher nicht gelungen ist, habe ich mir die Frage vorgelegt, ob es nicht wenigstens spezielle Klassen endlicher Gruppen gibt, bei denen eine einfachere Lösung des Problems der Darstellung durch Kollineationen möglich ist.

Hierzu gehören nun in erster Linie die Abelschen Gruppen endlicher Ordnung, für die im § 2 der vorliegenden Arbeit gezeigt werden soll, daß dieses Darstellungsproblem direkt in Angriff genommen werden kann, ohne die viel tiefer liegende allgemeine SCHURsche Theorie heranzuziehen, wie dies noch in [2] geschehen war; im Anschluß daran soll in § 3 die Frage nach der Darstellung Abelscher Gruppen endlicher Ordnung durch *reelle* Kollineationen behandelt werden.

Um endlich auch ein Beispiel für nicht-Abelsche Gruppen zu geben, die sich ohne Heranziehung der SCHURschen Darstellungstheorie einfach erledigen lassen, sollen in § 4 noch die Diedergruppen behandelt werden.

[1]) Hinweise in eckigen Klammern beziehen sich auf das Literaturverzeichnis am Ende dieser Arbeit.

[2]) Die nötigen Definitionen über Darstellungen durch Kollineationen sind im § 1 zusammengestellt.

§ 1. Definitionen.

Ist eine endliche Gruppe $\mathfrak{G}$ mit den Elementen E (= Einheitselement), $A, B, \ldots$ gegeben, so sprechen wir von einer *Darstellung* dieser Gruppe durch *Kollineationen* vom *Grade* f, wenn folgende Bedingungen erfüllt sind:

1. Jedem Element G von $\mathfrak{G}$ entspricht eine f-reihige quadratische Matrix $M(G)$ mit reellen oder komplexen Zahlen als Elementen und mit von Null verschiedener Determinante.

2. Für je zwei (nicht notwendig voneinander verschiedene) Elemente A, B aus $\mathfrak{G}$ existiert eine reelle oder komplexe Zahl $\lambda_{A,B}$ derart, daß

$$M(AB) = \lambda_{A,B} M(A) \cdot M(B) \tag{1.1}$$

gilt.

Der einzige Unterschied gegenüber der von FROBENIUS entwickelten klassischen Theorie der Darstellung durch ganze lineare homogene Substitutionen besteht in der Hinzufügung der Zahlenfaktoren $\lambda_{A,B}$ in (1.1); wir werden daher sagen, daß eine Darstellung durch Kollineationen gleichzeitig eine Darstellung durch ganze lineare homogene Substitutionen ist, wenn alle Zahlenfaktoren $\lambda_{A,B} = 1$ sind oder wenn durch Multiplikation der Matrizen $M(G)$ mit geeigneten Zahlenfaktoren λ_G erreicht werden kann, daß $\lambda_{A,B} = 1$ ausfällt für alle Elemente A, B von $\mathfrak{G}$.

In der Tat kann offenbar bei einer Darstellung durch Kollineationen jede Matrix $M(G)$ jederzeit mit einem ganz beliebigen Zahlenfaktor $\lambda_G \neq 0$ multipliziert werden (daher der Name Kollineationen!), weil dies nur eine unwesentliche Abänderung der Zahlenfaktoren $\lambda_{A,B}$ in (1.1) zur Folge hat. Wir nennen diese Operation eine „*Umnormierung*" und werden von dieser Möglichkeit mehrfach Gebrauch machen, da sie uns gestattet, die Matrizen einer Darstellung gegebenenfalls noch geeignet zu „normieren". So können wir z. B. durch Umnormierung stets erreichen, daß $M(E)$ gleich der f-reihigen Einheitsmatrix wird, die wir mit I_f bezeichnen wollen:

$$M(E) = I_f = \begin{pmatrix} 1 & 0 \ldots 0 \\ 0 & 1 \ldots 0 \\ \cdot & \cdot \quad \cdot \\ 0 & 0 \ldots 1 \end{pmatrix}, \tag{1.2}$$

was im folgenden immer vorausgesetzt werden soll. Zum Beweise brauchen wir nur festzustellen, daß sich aus (1.1) für $A = B = E$ die Tatsache ergibt, daß $M(E)$ der mit $1/\lambda_{E,E}$ multiplizierten f-reihigen Einheitsmatrix gleich sein muß.

Wie üblich nennen wir jede solche Matrix der Form λI_f (mit einem von Null verschiedenen Zahlenfaktor λ) eine *Multiplikation*; nach dem soeben Gesagten sind solche Multiplikationen bei Darstellungen durch Kollineationen als von der Einheitsmatrix nicht wesentlich verschieden anzusehen. Daher interessieren bei Kollineationen nur Darstellungen mit $f > 1$.

Wir werden eine Darstellung durch Kollineationen *getreu* nennen, wenn zwei verschiedenen Elementen von $\mathfrak{G}$ stets zwei wesentlich verschiedene Matrizen entsprechen (d.h. zwei Matrizen, die sich nicht nur durch einen Proportionalitätsfaktor unterscheiden); mit anderen Worten: eine Darstellung durch Kollineationen ist getreu, wenn eine Matrizengleichung

$$M(A) = \lambda M(B) \tag{1.3}$$

unmöglich ist für $A \neq B$. Ein häufig bequemeres Kriterium gibt der unmittelbar aus (1.1) und (1.2) folgende

Satz 1.1. *Eine Darstellung einer endlichen Gruppe durch Kollineationen ist dann und nur dann getreu, wenn keinem von E verschiedenen Gruppenelement eine Multiplikation zugeordnet ist.*

Die Äquivalenz zweier Darstellungen wird ähnlich wie in der klassischen Substitutionstheorie definiert: Zwei Darstellungen von $\mathfrak{G}$ durch Kollineationen, $M(G)$ und $N(G)$, heißen *äquivalent*, wenn es eine f-reihige Matrix P und Zahlen $\lambda_E, \lambda_A, \lambda_B, \ldots$ derart gibt, daß für alle Elemente G von $\mathfrak{G}$ die Matrizengleichung:

$$N(G) = \lambda_G \cdot P^{-1} \cdot M(G) \cdot P \tag{1.4}$$

erfüllt ist. (Der Unterschied gegenüber der Definition bei ganzen linearen homogenen Substitutionen besteht natürlich im Faktor λ_G auf der rechten Seite.) Äquivalente Darstellungen werden als nicht wesentlich verschieden betrachtet.

Wie in der klassischen Darstellungstheorie heißt auch eine Darstellung von $\mathfrak{G}$ durch Kollineationen *irreduzibel*, wenn es keine zu ihr äquivalente Darstellung $M(G)$ gibt, bei der alle Matrizen die Form

$$M(G) = \begin{pmatrix} M_1(G) & 0 \\ 0 & M_2(G) \end{pmatrix} \tag{1.5}$$

haben, wo $M_1(G)$ eine f_1-reihige quadratische Matrix und $M_2(G)$ eine f_2-reihige quadratische Matrix ist, während die beiden Nullen rechteckige, nur aus Nullen bestehende Matrizen andeuten. (f_1 und f_2 müssen vom betrachteten Gruppenelement G unabhängig sein!)

Endlich wollen wir noch eine Definition einführen, die sich bei SCHUR noch nicht findet, die sich aber als zweckmäßig erweist. Wir wollen eine Darstellung von $\mathfrak{G}$ durch Kollineationen $M(G)$ dann und nur dann eine *wesentliche* Darstellung nennen, wenn es keine reellen oder komplexen Zahlen $\lambda_E, \lambda_A, \lambda_B, \ldots$ derart gibt, daß die Matrizen $\lambda_G M(G)$ eine Darstellung von $\mathfrak{G}$ durch ganze lineare homogene Substitutionen liefern. Offenbar gilt dann:

Satz 1.2. *Eine Darstellung durch Kollineationen ist dann und nur dann wesentlich, wenn bei jeder Umnormierung der Matrizen $M(G)$ mindestens eine der durch (1.1) definierten Zahlen $\lambda_{A,B}$ von 1 verschieden ist.*

Gruppen, die keine wesentlichen irreduziblen Darstellungen durch Kollineationen zulassen, heißen bei SCHUR *abgeschlossen*.

§ 2. Allgemeine Theorie der Darstellung endlicher Abelscher Gruppen durch Kollineationen.

Ist $\mathfrak{A}$ eine Abelsche Gruppe endlicher Ordnung, so können wir sie stets als direktes Produkt zyklischer Gruppen geschrieben denken:

$$\mathfrak{A} = \{A_1\} \times \{A_2\} \times \cdots \times \{A_n\}, \tag{2.1}$$

wobei, damit n möglichst gering ausfällt, für die Ordnungen (oder Perioden) $a_1, a_2, \ldots, a_n$ der erzeugenden Elemente $A_1, A_2, \ldots, A_n$ angenommen werden kann, daß jede durch die folgende teilbar ist.

Die Gruppe (2.1) ist dann vollständig definiert durch die n Ordnungsrelationen:

$$A_i^{a_i} = E, \quad (i = 1, 2, \ldots, n) \tag{2.2}$$

und wenn $n > 1$ ist, durch die $\binom{n}{2}$ Vertauschbarkeitsrelationen:

$$A_j A_i = A_i A_j, \quad (1 \leq i < j \leq n) \tag{2.3}$$

wobei noch vorausgesetzt wird, daß die $n-1$ Quotienten:

$$\frac{a_i}{a_{i+1}} \quad (i = 1, 2, \ldots, n-1) \tag{2.4}$$

ganze Zahlen sein sollen.

Bei einer Darstellung von $\mathfrak{A}$ durch Kollineationen genügt es dann wegen (1.1), die den n erzeugenden Elementen A_i zugeordneten Matrizen $M(A_i)$ $(i = 1, 2, \ldots, n)$ anzugeben, denn durch Umnormierung können wir stets erreichen, daß irgendeinem Gruppenelement $A_1^{\alpha_1} A_2^{\alpha_2} \ldots A_n^{\alpha_n}$ die Matrix:

$$M(A_1^{\alpha_1} A_2^{\alpha_2} \ldots A_n^{\alpha_n}) = \{M(A_1)\}^{\alpha_1} \cdot \{M(A_2)\}^{\alpha_2} \cdot \ldots \cdot \{M(A_n)\}^{\alpha_n} \quad (0 \leq \alpha_i < a_i) \tag{2.5}$$

entsprechen möge.

Wegen (1.1) und (2.2) müssen die Matrizen $\{M(A_i)\}^{a_i}$ Multiplikationen sein; durch Umnormierung können wir sogar erreichen, daß:

$$\{M(A_i)\}^{a_i} = I_f \quad (i = 1, 2, \ldots, n) \tag{2.6}$$

wird. Diese Normierung hat übrigens zur Folge, daß

$$M(A_i^{-1}) = \{M(A_i)\}^{-1} \quad (i = 1, 2, \ldots, n) \tag{2.7}$$

wird; denn wegen (2.2) und (2.5) wird

$$M(A_i^{-1}) = M(A_i^{a_i - 1}) = \{M(A_i)\}^{a_i - 1} = \{M(A_i)\}^{a_i} \cdot \{M(A_i)\}^{-1} = \{M(A_i)\}^{-1}.$$

Satz 2.1. *Zyklische Gruppen besitzen überhaupt keine wesentlichen Darstellungen durch Kollineationen*[3]).

In der Tat ergibt sich aus (2.6) im Falle $n = 1$, daß bei der von uns vorgenommenen Normierung jede Darstellung durch Kollineationen in eine Darstellung durch ganze lineare homogene Substitutionen übergeht.

[3]) In der SCHURschen Terminologie lautet Satz 2.1: „Zyklische Gruppen sind abgeschlossen." Er findet sich in dieser Formulierung in [3], S. 50.

Da durch Satz 2.1 also das Darstellungsproblem für zyklische Gruppen völlig auf das der Darstellung durch ganze lineare Substitutionen zurückgeführt ist, können wir im folgenden $n \geq 2$ voraussetzen und dann folgenden Satz aussprechen:

Satz 2.2. *Dafür, daß irgendwelche n Matrizen* $M(A_1), M(A_2), \ldots, M(A_n)$, *die den Bedingungen* (2.6) *genügen, auf Grund von* (2.5) *eine Darstellung von* $\mathfrak{A}$ *durch Kollineationen liefern, ist notwendig und hinreichend die Existenz von* $\binom{n}{2}$ *reellen oder komplexen Zahlen* λ_{ij} $(1 \leq i < j \leq n)$, *mit denen die folgenden Matrizengleichungen erfüllt sind*:

$$M(A_i) \cdot M(A_j) = \lambda_{ij} M(A_j) \cdot M(A_i) \qquad (1 \leq i < j \leq n). \tag{2.8}$$

In der Tat erfüllen die Matrizen $M(A_i)$ wegen (2.6) die Ordnungsrelationen (2.2) und wegen (2.8) — bis auf laut (1.1) zulässige Zahlenfaktoren λ_{ij} — die Vertauschbarkeitsrelationen (2.3).

Satz 2.3. λ_{ij} *muß eine* a_j*-te Einheitswurzel sein; bei einer wesentlichen Darstellung durch Kollineationen dürfen nicht alle* $\lambda_{ij} = 1$ *sein.*

Beweis. Schreiben wir (2.8) in der Form:

$$M(A_j) = \lambda_{ij} \{M(A_i)\}^{-1} \cdot M(A_j) \cdot M(A_i) \qquad (1 \leq i < j \leq n) \tag{2.9}$$

und erheben wir beide Seiten in die a_j-te Potenz, so ergibt sich wegen (2.6):

$$I_f = \lambda_{ij}^{a_j} \{M(A_i)\}^{-1} \cdot I_f \cdot M(A_i) = \lambda_{ij}^{a_j} \cdot I_f$$

und daher:

$$\lambda_{ij}^{a_j} = 1, \qquad (1 \leq i < j \leq n) \tag{2.10}$$

womit der erste Teil von Satz 2.3 bewiesen ist; der zweite Teil ist eine unmittelbare Folge von (2.8), Satz 1.2 und der Tatsache, daß die λ_{ij} sich nicht ändern, auch wenn wir die Matrizen $M(A_i)$ zwecks Umnormierung mit irgendwelchen Zahlenfaktoren multiplizieren.

λ_{ij} muß übrigens auch eine f-te Einheitswurzel sein, denn wenn wir auf beiden Seiten von (2.9) die Determinanten nehmen und durch die von 0 verschiedene Determinante der Matrix $M(A_j)$ dividieren, ergibt sich:

$$\lambda_{ij}^{f} = 1 \qquad (1 \leq i < j \leq n). \tag{2.11}$$

Schreiben wir endlich (2.8) in der Form:

$$\{M(A_i)\}^{-1} \cdot \{M(A_j)\}^{-1} \cdot M(A_i) \cdot M(A_j) = \lambda_{ij} \cdot I_f, \tag{2.12}$$

so können wir Satz 2.2 auch folgendermaßen formulieren:

Satz 2.4. *n den Bedingungen* (2.6) *genügende Matrizen* $M(A_1), M(A_2), \ldots, M(A_n)$ *liefern dann und nur dann eine Darstellung von* $\mathfrak{A}$ *durch Kollineationen, wenn die* $\binom{n}{2}$ *„Kommutatoren"* $\{M(A_i)\}^{-1} \cdot \{M(A_j)\}^{-1} \cdot M(A_i) \cdot M(A_j)$ *Multiplikationen sind* $(1 \leq i < j \leq n)$[4].

[4]) Bei einer *wesentlichen* Darstellung durch Kollineationen muß natürlich mindestens einer dieser Kommutatoren von der Einheitsmatrix I_f verschieden sein.

Andererseits sind nach einem klassischen Satz der Darstellungstheorie die einzigen Matrizen, die mit sämtlichen Matrizen einer *irreduziblen* Substitutionsgruppe vertauschbar sind, Multiplikationen. Es liegt daher nahe, unser Darstellungsproblem dadurch zu lösen, daß man die irreduziblen Darstellungen durch ganze lineare Substitutionen einer abstrakten Gruppe $\mathfrak{H} = \{H_1, H_2, \ldots, H_n\}$ betrachtet, die folgendermaßen zu definieren ist: Die erzeugenden Elemente H_i sollen den (2.2) analogen Ordnungsrelationen

$$H_i^{a_i} = E \qquad (i = 1, 2, \ldots, n) \tag{2.13}$$

und den (2.3) entsprechenden, aber schwächeren Vertauschbarkeitsbedingungen genügen, daß die $\binom{n}{2}$ Kommutatoren

$$C_{ij} = H_i^{-1} H_j^{-1} H_i H_j \qquad (1 \leq i < j \leq n) \tag{2.14}$$

mit allen Elementen von $\mathfrak{H}$ vertauschbar sein sollen; es genügt natürlich zu verlangen, daß die $n^2(n-1)/2$ Vertauschbarkeitsrelationen:

$$H_i C_{jk} = C_{jk} H_i \qquad (j < k;\ i, j, k = 1, 2, \ldots, n) \tag{2.15}$$

erfüllt sein sollen.

Genau wie in [2], S. 19, beweist man dann leicht (ohne Heranziehung der SCHURschen Darstellungstheorie), daß $\mathfrak{H}$ eine endliche Gruppe der Ordnung $a_1 a_2^2 a_3^3 \ldots a_n^n$ ist.

Wenn nun eine *irreduzible* Darstellung von $\mathfrak{H}$ durch ganze lineare homogene Substitutionen vorliegt, bei der dem Element H von $\mathfrak{H}$ die Matrix $N(H)$ entsprechen möge, so können wir

$$M(A_i) = N(H_i) \qquad (i = 1, 2, \ldots, n) \tag{2.16}$$

setzen und *erhalten auf diese Weise eine irreduzible Darstellung von $\mathfrak{A}$ durch Kollineationen*, denn wegen (2.13) ist (2.6) erfüllt, und da definitionsgemäß die C_{ij} invariante Elemente von $\mathfrak{H}$ sind, sind auf Grund des oben zitierten Satzes die Matrizen

$$\left\{ \begin{aligned} N(C_{ij}) &= \{N(H_i)\}^{-1} \{N(H_j)\}^{-1} N(H_i)\, N(H_j) \\ &= \{M(A_i)\}^{-1} \{M(A_j)\}^{-1} M(A_i)\, M(A_j) \end{aligned} \right. \tag{2.17}$$

Multiplikationen und somit alle im Satz 2.4 ausgesprochenen Bedingungen erfüllt. Dieselbe Schlußweise läßt sich anwenden bei einer nicht notwendigerweise irreduziblen Darstellung $N(H)$ von $\mathfrak{H}$ durch ganze lineare homogene Substitutionen, sofern nur bei derselben die $N(C_{ij})$ Multiplikationen sind.

Ist umgekehrt eine (nicht notwendig irreduzible) Darstellung von $\mathfrak{A}$ durch Kollineationen gegeben, so erfüllen die durch (2.6) normierten Matrizen $M(A_i)$ die Ordnungsrelationen (2.13) für die H_i und als Multiplikationen erfüllen die Kommutatoren (2.12) gewiß die an die Kommutatoren (2.14) gestellten Bedingungen (2.15).

Zusammenfassend können wir daher sagen:

Satz 2.5. *Die Darstellungen von $\mathfrak{A}$ durch Kollineationen können erhalten werden, indem man diejenigen Darstellungen der durch* (2.13) *und* (2.15) *definierten abstrakten Gruppe $\mathfrak{H}$ durch ganze lineare homogene Substitutionen aufstellt, bei denen die Kommutatoren* (2.14) *durch Multiplikationen dargestellt werden, wie dies insbesondere bei irreduziblen Darstellungen von $\mathfrak{H}$ der Fall ist, denen natürlich die irreduziblen Darstellungen von $\mathfrak{A}$ durch Kollineationen entsprechen.*

Damit haben wir aber für den hier betrachteten Fall Abelscher Gruppen durch ganz elementare Betrachtungen dasselbe Ergebnis erhalten, das für beliebige endliche Gruppen das Ziel der viel tiefer liegenden SCHURschen Darstellungstheorie ist: die irreduziblen Darstellungen durch Kollineationen dadurch zu erhalten, daß man die irreduziblen Darstellungen einer größeren Gruppe, der „Darstellungsgruppe" in der SCHURschen Terminologie, durch ganze lineare homogene Substitutionen betrachtet. In der Tat ist (abgesehen von der Bezeichnungsweise der erzeugenden Elemente) die durch (2.13) und (2.15) definierte Gruppe $\mathfrak{H}$ mit der ebenfalls $\mathfrak{H}$ heißenden „einfachsten Darstellungsgruppe" von [2], S. 19 identisch, die dort unter Anwendung der SCHURschen Darstellungstheorie erhalten wurde.

Da dann in den §§ 2 bis 6 von [2] die SCHURsche Theorie selbst nicht mehr herangezogen wird, können wir die weiteren Ausführungen von dort unverändert übernehmen und auf eine Wiederholung der dortigen Untersuchungen und Ergebnisse hier verzichten[5]); statt dessen soll im folgenden § 3 lieber die in [2] nicht behandelte Frage untersucht werden, welche Abelsche Gruppen wesentliche Darstellungen durch *reelle* Kollineationen zulassen.

§ 3. Bedingungen für die Existenz wesentlicher getreuer Darstellungen einer Abelschen Gruppe durch reelle Kollineationen.

Wir wollen nunmehr nach *reellen* Kollineationen fragen, d.h. nach solchen Darstellungen, wo sowohl die Matrizen $M(G)$ nur reelle Elemente besitzen als auch infolgedessen die durch (1.1) definierten Zahlen $\lambda_{A,B}$ alle reell sind.

Jede endliche Gruppe besitzt reelle getreue Darstellungen durch ganze lineare homogene Substitutionen (und daher auch durch Kollineationen); z.B. braucht man nur die der Gruppentafel entsprechende Darstellung durch reguläre Permutationen in Matrizenform zu schreiben.

Anders wird das, wenn wir nach *wesentlichen* reellen getreuen Darstellungen durch Kollineationen fragen, die also nicht Darstellungen durch ganze lineare

[5]) Es soll nur bei dieser Gelegenheit darauf hingewiesen werden, daß § 5 von [2] leider einen Irrtum enthält; entgegen der in Anm. 26 aufgestellten Behauptung braucht (48) nicht notwendigerweise ein Automorphismus zu sein. Trotzdem gelingt es auch dann — freilich manchmal nur durch viel längere Rechnungen — durch ein dem in [2] geschilderten ähnliches Kompositionsverfahren alle irreduziblen getreuen Darstellungen zu finden.

homogene Substitutionen sind oder durch Umnormierung sich in solche überführen lassen. Wir sahen bereits oben (Satz 2.1), daß zyklische Gruppen überhaupt keine wesentlichen Darstellungen durch Kollineationen zulassen, geschweige denn reelle. Wie liegt die Sache bei nichtzyklischen Abelschen Gruppen? Die Antwort wird gegeben durch

Satz 3.1. *Dann und nur dann, wenn $n \geq 2$ ist und wenn außerdem a_2 eine gerade Zahl ist, besitzt die Gruppe* (2.1) *wesentliche getreue Darstellungen durch reelle Kollineationen.*

Beweis. Daß $n \geq 2$ sein muß, folgt aus Satz 2.1; die Notwendigkeit der Bedingung, daß a_2 (und mithin auch a_1) eine gerade Zahl sein muß, erkennt man folgendermaßen indirekt:

Wäre a_2 eine ungerade Zahl, so wären wegen (2.4) auch $a_3, a_4, \ldots, a_n$ ungerade Zahlen. Die einzige *reelle* Lösung der Gln. (2.10) wäre dann: $\lambda_{ij} = 1$ für alle $1 \leq i < j \leq n$; das ist aber bei einer wesentlichen Darstellung nach Satz 2.3 unmöglich.

Daß die im Satz 3.1 angegebenen Bedingungen für die Existenz reeller getreuer wesentlicher Darstellungen durch Kollineationen auch hinreichend sind, beweist man durch Angabe eines Beispiels: Wie jede Gruppe läßt auch unsere Abelsche Gruppe $\mathfrak{A}$ eine reelle getreue Darstellung durch ganze lineare homogene Substitutionen zu, z. B. die in Matrizenform geschriebene Darstellung durch reguläre Permutationen. Sei hierbei $P(A_i)$ die dem erzeugenden Element A_i $(i = 1, 2, \ldots, n)$ entsprechende Matrix[6]). Wir bilden dann folgende quadratische Matrizen mit $2a_1a_2 \ldots a_n$ Zeilen und Spalten:

$$M(A_1) = \begin{pmatrix} P(A_1) & 0 \\ 0 & -P(A_1) \end{pmatrix}, \tag{3.1}$$

$$M(A_2) = \begin{pmatrix} 0 & P(A_2) \\ P(A_2) & 0 \end{pmatrix}, \tag{3.2}$$

und wenn $n > 2$ ist:

$$M(A_i) = \begin{pmatrix} P(A_i) & 0 \\ 0 & P(A_i) \end{pmatrix} \qquad (i = 3, 4, \ldots, n). \tag{3.3}$$

Daß es sich hierbei um eine reelle getreue Darstellung von $\mathfrak{A}$ durch Kollineationen handelt, wenn a_2 gerade ist, prüft man leicht nach; die Darstellung ist wesentlich, weil $\lambda_{12} = -1$ ist.

Hiermit ist der Beweis von Satz 3.1 beendet; es sei jedoch bemerkt, daß man statt der hier verwendeten Darstellung (3.1) bis (3.3) reelle Darstellungen von geringerem Grad finden kann, indem man für die $P(A_i)$ statt der regulären Permutationen eine andere reelle getreue Darstellung von $\mathfrak{A}$ durch ganze lineare homogene Substitutionen einsetzt. Manchmal ist noch eine weitere Reduktion des Grades möglich, wie aus dem folgenden Beispiel ersichtlich

[6]) So daß $\{P(A_i)\}^{a_i} = I_{\Pi a_i}$ und $P(A_i) \cdot P(A_j) = P(A_j) \cdot P(A_i)$ gilt.

ist, wo für die Abelsche Gruppe der Ordnung 12 mit $a_1 = 6$, $a_2 = 2$ eine wesentliche getreue Darstellung durch reelle Kollineationen vom vierten Grade (statt vom Grade $2a_1a_2 = 24$) angegeben wird:

$$M(A_1) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 & 0 \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ 0 & 0 & \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \tag{3.4}$$

$$M(A_2) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{3.5}$$

§ 4. Darstellung von Diedergruppen durch Kollineationen.

$\mathfrak{D}_{2n}$, die Diedergruppe der Ordnung $2n$ (wo $n > 2$ vorausgesetzt ist), kann bekanntlich durch zwei Elemente A und B mit den folgenden definierenden Relationen erzeugt werden:

$$A^2 = B^2 = (AB)^n = E. \tag{4.1}$$

Da die $2n$ Elemente von $\mathfrak{D}_{2n}$ in der Form $(AB)^\alpha \cdot B^\beta$ $(0 \leq \alpha < n;\ 0 \leq \beta \leq 1)$ geschrieben werden können, kann bei einer Darstellung von $\mathfrak{D}_{2n}$ durch Kollineationen durch passende Normierung stets erreicht werden, daß

$$M\big((AB)^\alpha \cdot B^\beta\big) = \{M(AB)\}^\alpha \cdot \{M(B)\}^\beta \quad (\alpha = 0, 1, \ldots, n-1;\ \beta = 0, 1) \tag{4.2}$$

wird; es genügt daher, die beiden Matrizen $M(AB)$ und $M(B)$ anzugeben oder statt dessen auch die Matrizen $M(A)$ und $M(B)$, denn aus (4.2) ergibt sich insbesondere für $\alpha = \beta = 1$:

$$M(A) = M(AB) \cdot M(B). \tag{4.3}$$

Ferner müssen wegen (4.1) die beiden Matrizen $\{M(A)\}^2$ und $\{M(B)\}^2$ Multiplikationen sein; durch Umnormierung können wir daher stets erreichen, daß

$$\{M(A)\}^2 = \{M(B)\}^2 = I_f \tag{4.4}$$

wird, und durch rechtsseitige Multiplikation von (4.3) mit $M(B)$ ergibt sich dann:

$$M(AB) = M(A) \cdot M(B). \tag{4.5}$$

Endlich muß wegen (4.1) auch $\{M(AB)\}^n$ eine Multiplikation sein; d. h. es muß eine reelle oder komplexe Zahl λ geben derart, daß

$$\{M(AB)\}^n = \{M(A) \cdot M(B)\}^n = \lambda I_f \tag{4.6}$$

gilt. Hier läßt sich aber zeigen, daß $\lambda^2 = 1$ sein muß:

Multiplizieren wir (4.3) links mit $M(B)$, so ergibt sich wegen (4.4):

$$M(B) \cdot M(A) = \{M(B)\}^{-1} \cdot M(AB) \cdot M(B); \tag{4.7}$$

wegen (4.4) und (4.5) ist aber die linke Seite von (4.7) gleich $\{M(AB)\}^{-1}$, so daß wir erhalten:

$$\{M(AB)\}^{-1} = \{M(B)\}^{-1} \cdot M(AB) \cdot M(B). \tag{4.8}$$

Erheben wir nun beide Seiten von (4.8) in die n-te Potenz, so folgt:

$$\{M(AB)\}^{-n} = \{M(B)\}^{-1} \cdot \{M(AB)\}^{n} \cdot M(B) \tag{4.9}$$

und durch Einsetzen von (4.6) ergibt sich

$$\frac{1}{\lambda} I_f = \lambda I_f, \tag{4.10}$$

d.h. tatsächlich:

$$\lambda = \pm 1. \tag{4.11}$$

Bei einer *wesentlichen* Darstellung von $\mathfrak{D}_{2n}$ durch Kollineationen muß aber dann

$$\lambda = -1 \tag{4.12}$$

sein (da bei den vorgenommenen Normierungen eine Darstellung durch ganze lineare homogene Substitutionen vorliegen würde, wenn $\lambda = 1$ wäre).

Bei einer wesentlichen Darstellung geht daher (4.6) über in:

$$\{M(AB)\}^{n} = \{M(A) \cdot M(B)\}^{n} = -I_f. \tag{4.13}$$

Wir müssen nun folgende zwei Fälle unterscheiden:

1. *n ist ungerade*; dann können wir, ohne die Normierungsbedingungen (4.4) zu verletzen, die Matrix $M(A)$ noch mit -1 multiplizieren, die Matrix $M(B)$ aber unverändert lassen, so daß sich laut (4.5) auch $M(AB)$ mit -1 multipliziert. Bei dieser Umnormierung verschwindet aber das Minuszeichen in (4.13); es gibt also in diesem Falle gar keine wesentlichen Darstellungen durch Kollineationen.

2. *n ist gerade*; jetzt kann das Minuszeichen in (4.13) nicht durch Multiplikation von $M(A)$ oder $M(B)$ mit -1 weggeschafft werden. Tatsächlich gibt es jetzt wesentliche Darstellungen durch Kollineationen, wie die folgende Darstellung $N(G)$ lehrt, in der ε eine primitive $2n$-te Einheitswurzel bedeutet:

$$N(A) = \begin{pmatrix} 0 & \varepsilon \\ 1 & 0 \\ \varepsilon & \end{pmatrix}, \quad N(B) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.14}$$

Es ist aber auch leicht einzusehen, wie man bei geradem n alle wesentlichen Darstellungen erhalten kann. Erhebt man nämlich (4.13) ins Quadrat, so ergibt sich:

$$\{M(A) \cdot M(B)\}^{2n} = I_f \tag{4.15}$$

und somit erfüllen wegen (4.4) die Matrizen $M(A)$ und $M(B)$ die zu (4.1) ganz analogen definierenden Relationen einer Diedergruppe der doppelten

Ordnung $4n$. Unter Berücksichtigung von (4.13) können wir also endlich folgenden Satz[7]) aussprechen:

Satz 4.1. *Die Diedergruppe der Ordnung $2n$ besitzt dann und nur dann wesentliche Darstellungen durch Kollineationen, wenn n gerade ist; diese können bei geeigneter Normierung erhalten werden als solche Darstellungen einer Diedergruppe der Ordnung $4n$ durch ganze lineare homogene Substitutionen, bei denen $(AB)^n$ durch die mit -1 multiplizierte Einheitsmatrix dargestellt wird.*

Literatur.

[1] Asano, K., u. K. Shoda: Zur Theorie der Darstellungen einer endlichen Gruppe durch Kollineationen. Compositio Mathematica **2**, 230—240 (1935). — [2] Frucht, R.: Über die Darstellung endlicher Abelscher Gruppen durch Kollineationen. J. reine u. angew. Math. **166**, 16—29 (1931). — [3] Schur, I.: Über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **127**, 20—50 (1904). — [4] Schur, I.: Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **132**, 85—137 (1907). — [5] Schur, I.: Über die Darstellung der symmetrischen und alternierenden Gruppe durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **139**, 155—250 (1911).

[7]) In der Schurschen Terminologie würde dieser Satz folgendermaßen lauten: „$\mathfrak{D}_{2n}$ ist abgeschlossen, wenn n ungerade ist; wenn n gerade ist, ist $\mathfrak{D}_{4n}$ eine Darstellungsgruppe von $\mathfrak{D}_{2n}$." Bei Schur findet sich dieser Satz in [4], S. 108f., aber nur für den Fall, daß n eine Potenz von 2 ist; es ist nämlich zu beachten, daß die von Schur dort $\mathfrak{D}_l$ genannte Gruppe die Diedergruppe der Ordnung 2^l ist (also unser $\mathfrak{D}_{2^l}$). Ist n eine ungerade *quadratfreie* Zahl, so ist die Tatsache, daß $\mathfrak{D}_{2n}$ abgeschlossen ist, auch eine Folge eines allgemeinen Satzes von Schur, wonach alle Gruppen von quadratfreier Ordnung abgeschlossen sind (s. [3], S. 48).

Universidad Técnica Santa Maria, Cas. 110-V, Valparaiso (Chile).

(Eingegangen am 30. Januar 1955.)

Hopf, E.
Math. Zeitschr. Bd. 63, S. 156—157 (1955)

Die Harnacksche Ungleichung für positive harmonische Funktionen.

Issai Schur in memoriam.

Von

Eberhard Hopf.

G bezeichne ein offenes Gebiet im n-dimensionalen Euklidischen Raume. Ein bekannter Satz der Potentialtheorie besagt, daß für irgend zwei feste Punkte x, x' in G die für sämtliche in G positiven und harmonischen Funktionen u betrachteten Quotienten $u(x')/u(x)$ eine endliche obere Grenze besitzen. Ebenfalls bekannt ist, daß für dieselben Funktionen u der logarithmische Gradient $u'(x)/u(x)$ in jedem festen Punkte x von G beschränkt ist. Die letztere Tatsache gehört von Rechts wegen in dieselbe Kategorie wie einige neuerdings entdeckte interessante Ungleichungen für Ableitungen von Lösungen gewisser nichtlinearer elliptischer Differentialgleichungen (z. B. der Satz von E. Heinz über die zweiten Ableitungen der Lösungen der Minimalflächengleichung). Man sieht das, wenn man bedenkt, daß Positivität und Harmonizität von u in G zusammen genau bedeuten, daß $v = \ln u$ eine in G reguläre Lösung der nichtlinearen Differentialgleichung

$$\Delta v + \sum \left(\frac{\partial v}{\partial x_\nu} \right)^2 = 0$$

ist. Jene Tatsache besagt dann, daß für alle in G regulären Lösungen v die ersten Ableitungen $v'(x)$ in jedem festen Punkte x von G beschränkt sind. Diese Bemerkung sei nur nebenbei gemacht. Der Hauptzweck dieser Note ist, einen besonders einfachen Beweis der beiden Sätze über positive harmonische Funktionen mitzuteilen. Die Sätze haben ihre Wurzel in ihrem wichtigsten Spezialfall, wo G ein Hyperkugelgebiet ist (Harnacksche Ungleichungen). Es genügt, den Fall zu betrachten, wo G das Kugelgebiet $|x| < R$ um den Nullpunkt ist, und den Quotienten $u(x)/u(0)$ nach oben abzuschätzen. Die übliche Herleitung dieser Abschätzung beruht auf der Poissonschen Integralformel, die die erste Randwertaufgabe für den Kugelfall explizit löst. Man kann jedoch schon ganz einfach schließen, wenn man lediglich die Tatsache benützt, daß für eine in einem abgeschlossenen Kugelbereich harmonische Funktion der Wert im Mittelpunkt gleich dem Integralmittel der Werte auf der Oberfläche ist. Daraus folgt unmittelbar, daß der Wert im Mittelpunkt auch gleich dem Volummittel der Werte in der Kugel ist, und das bleibt richtig, wenn man bei der Mittelbildung einen beliebigen Gewichtsfaktor benützt, vorausgesetzt, daß dieser nur von der Entfernung vom Mittelpunkt abhängt. Zum Beispiel ist der Wert im Mittelpunkt gleich dem gewöhnlichen Volummittel der Werte in einer konzentrischen Kugelschale.

u sei harmonisch und positiv in $K: |y| \leq R$. $V(r) \sim r^n$ bezeichne das Volumen einer Kugel vom Radius r. Die größte in K gelegene Kugel K' um einen Punkt x von K hat den Radius $R-|x|$. Dann folgt wegen $u>0$

$$V(R)\,u(0) = \int_K u(y)\,dV \geq \int_{K'} u(y)\,dV = V(R-|x|)\,u(x)$$

oder

$$\frac{u(x)}{u(0)} \leq \left(\frac{R}{R-|x|}\right)^n, \tag{1}$$

womit bereits alles bewiesen ist. Subtrahiert man Eins von beiden Seiten und läßt man nach Division durch $|x|$ diese letztere Größe gegen Null streben, so erhält man

$$\frac{|u'(0)|}{u(0)} \leq \frac{n}{R} \tag{2}$$

für den logarithmischen Gradienten im Nullpunkt. (1) und (2) gelten, wenn u lediglich im Inneren von K, $|x|<R$, harmonisch und positiv ist (triviale Grenzbetrachtung). Aus (2) folgt sofort die allgemeinere Ungleichung: Ist u in einem offenen Gebiete G harmonisch und positiv, so gilt in G

$$\frac{|u'(x)|}{u(x)} \leq \frac{n}{d}, \tag{2'}$$

wo d der Abstand von x vom Rande von G ist. Wendet man (2') wieder auf die Kugel K an, so ergibt sich durch Integration längs eines Radius von 0 bis x wieder genau die Ungleichung (1).

Die Schranke in (2) ist bekanntlich die beste. In (1) ist das jedoch nicht der Fall. Da liefert die Poissonsche Integralformel die richtige obere Grenze. Die letztere wird am Rande $|x|=R$ nicht von der n-ten, sondern nur von der $(n-1)$-ten Ordnung unendlich. Die richtige Größenordnung ergibt sich, wenn man den obigen Beweis von (1) nicht mit dem Volummittel über die ganze Kugel K, sondern mit dem Mittel über die dünnste, K' enthaltende Kugelschale S um 0 beginnt. S hat das Volumen $(|x|>\frac{1}{2}R)\ V(R)-V(2|x|-R)$. Damit erhält man

$$\frac{u(x)}{u(0)} \leq \frac{R^n-(2|x|-R)^n}{(R-|x|)^n}, \qquad |x| \geq \frac{1}{2}R.$$

Eine weitere Verbesserung der Schranke dürfte durch andere Gewichtsverteilung bei der Mittelbildung möglich sein, aber damit wäre die Einfachheit des Gedankenganges zerstört.

Grad. Inst. for Math. and Mech., University of Indiana, Bloomington, Indiana (USA.).

(Eingegangen am 15. Januar 1955.)

FEKETE, M., and G. SZEGÖ
Math. Zeitsohr. Bd. 63, S. 158—172 (1955)

On algebraic equations with integral coefficients whose roots belong to a given point set.

To the memory of I. SCHUR.

By
M. FEKETE and G. SZEGÖ.

§ 1. Introduction.

1. Generalizing certain results of I. SCHUR [6][1]) published in the first volume of this journal, M. FEKETE [2, pp. 240–241] proved the following theorem:

A. *Let S be a bounded and closed set in the complex z-plane, whose transfinite diameter d satisfies the condition*

$$d < 1. \tag{1}$$

Then there is only a finite number of irreducible algebraic equations with rational integral coefficients of the form

$$z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0 \tag{2}$$

the roots of which lie all in S.

For the definition of the transfinite diameter of a closed set, we refer to the above quoted paper of M. FEKETE. The condition (1) can be replaced, as mentioned there, by the condition

$$d^* < 1$$

where d^* denotes the transfinite diameter[2]) of the socalled "symmetric kernel" S^* of S; S^* is the set of all those points of S which belong, together with their images with respect to the real axis, to S.

2. By retaining condition (1), one can obtain another improvement of Theorem A, admitting algebraic equations of the form (2) whose coefficients are integral numbers of a given quadratic-imaginary field, and which are irreducible in this field. Thus the following theorem holds:

B. *Let S be a point-set as in Theorem A. Furthermore, let ϑ satisfy a quadratic equation*

$$z^2 + pz + q = 0$$

[1]) Numbers in square brackets refer to the Bibliography.

[2]) From the definition of the transfinite diameter it follows that $d^* \leq d$. If S^* is empty we write $d^* = 0$.

with rational integral coefficients and of a negative discriminant. Then there exists only a finite number of equations of the form (2), *irreducible in the algebraic field* $P(\vartheta)$ *generated by* ϑ, *whose coefficients are integers of* $P(\vartheta)$ *and whose roots lie all in* S.

In § 2 we shall give a proof of this theorem based on the "resultant method" of the paper [2].

3. Theorem B loses its validity when the algebraic field F to which the integral coefficients of (2) belong and in which this equation is irreducible, is different from that of the rational numbers or from the quadratic-imaginary fields $P(\vartheta)$. In fact, the integral numbers of such a field F lie every-where dense on the real axis or in the complex plane according as F is real or complex. Consequently an arbitrary small real interval or an arbitrary small circle contains an infinite number of algebraic integers from F. Thus there is an *infinite number* of equations of the form (2) with integral coefficients from F (for instance, linear ones) whose roots belong to a suitable chosen set S of an arbitrary small transfinite diameter. We can choose S so that it contains an arbitrary small real interval, or an arbitrary small circle.

4. We introduce the following terminology used throughout this paper. Let ϑ be a fixed number satisfying an algebraic equation $z^2+pz+q=0$ with rational integral p and q and of a negative discriminant, and let $P(\vartheta)$ denote the algebraic field generated by ϑ. We say that an algebraic equation is of the class $K(\vartheta)$ if it has the form (2) with integral coefficients belonging to $P(\vartheta)$ and it is irreducible in $P(\vartheta)$. Likewise, we say that an equation is of the class K if it has the form (2) with rational integral coefficients and it is irreducible in the rational field P.

From Theorem B it is obvious that the existence of an infinite number of equations of class $K(\vartheta)$ whose roots lie all in a closed and bounded set S, imposes upon the transfinite diameter d of S the inequality

$$d \geq 1.$$

Similarly the symmetric kernel S^* of S must have (see above) a transfinite diameter d^*, satisfying

$$d^* \geq 1$$

provided the equations (2) have rational integral coefficients only. The remark of 3. suggests, however, that these necessary conditions are not always sufficient in order to ensure the mentioned existence, but that the set S is to be subject to some further conditions as to the manner of the distribution of its points.

In fact, let S consist of the points of the circle $|z|=e$, or, more generally, of those of the circle $|z|=t$ where $t>1$ is a transcendental number; then —though $d(S)=d(S^*)=t>1$—there are no irreducible equations (2) with integral coefficients from a quadratic-imaginary field $P(\vartheta)$, whose roots lie all in S. Indeed this would imply the relation $t^n=|a_n|$ which is impossible.

5. It is an important but seemingly a difficult problem to find simple necessary and sufficient conditions for the just discussed existence. Instead of solving this problem we shall be satisfied in the present paper with the statement of some sufficient conditions, which, nevertheless, will lead us to an answer in some important and comparatively general cases. The conditions are of set-theoretic nature and will be formulated in § 4 and § 5.

We have already specified the content of § 2. In § 3 we deal with our problem for the case of lemniscates, respectively sets S which enclose a lemniscate-line of a certain type (Theorems C, D and E). The transition to more general sets follows in § 4 (Theorems F and G).

Subsequently, we derive in § 5 extensions of the results so obtained (Theorems H, I, J, K and L).

The proofs are of elementary nature. The decisive question is the possibility of the exhaustion of a given set through lemniscate-domains lying in this set and having transfinite diameters which converge to that of the set (respectively to the transfinite diameter of the set which results through adjunction of the boundary points of an open set). In all such cases the condition $d > 1$ ensures the existence of an infinite number of algebraic equations with integral coefficients from an arbitrarily prescribed imaginary-quadratic field $P(\vartheta)$, which are irreducible in $P(\vartheta)$ and whose roots belong to the given set.

§ 2. Proof of Theorem B.

1. By assumption (1) there must be [2, p. 236] a polynomial

$$b(z) = z^k + b_1 z^{k-1} + \cdots + b_{k-1} z + b_k$$

whose maximum-modulus μ in S is less than 1; consequently the polynomials

$$t_{km+l}(z) = [b(z)]^m z^l = z^{km+l} + \cdots, \qquad m = 0, 1, 2, \ldots; \; l = 0, 1, \ldots, k-1,$$

satisfy in S the inequality

$$|t_{km+l}(z)| \leq \gamma \varrho^{km+l} \tag{3}$$

where $\varrho = \mu^{1/k}$ and $\gamma \geq 1$ denote certain constants independent of m and l. Hence we can, as it will be shown below, infer the existence of a polynomial

$$c(z) = z^r + c_1 z^{r-1} + \cdots + c_{r-1} z + c_r$$

whose coefficients are integers of $P(\vartheta)$ (even of the form $g + h\vartheta$ with rational integral g and h) and whose absolute value on S remains under 1.

Now let

$$a(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$$

be any polynomial in z of the class $K(\vartheta)$ whose zeros $z_1, z_2, \ldots, z_n$ lie all in S; then the absolute value of the product

$$c(z_1)\, c(z_2) \ldots c(z_n) = \operatorname{Res}[c(z), a(z)] = \text{an integral number of } P(\vartheta)$$

must turn out to be less than 1; consequently, the resultant of the polynomials $a(z)$ and $c(z)$ is equal to zero. (Indeed, [1, p. 139] no integer of a quadratic-imaginary field $P(\vartheta)$ which is different from zero can have absolute value less than 1.) This last result means [with regard to the irreducibility of $a(z)$] that $a(z)$ must be a divisor of $c(z)$. But $c(z)$ can only have a finite number of different divisors with integral coefficients from $P(\vartheta)$. Therefore, the number of the polynomials $a(z)$ must be also finite, as asserted before.

2. The existence of the polynomial $c(z)$ with the desired properties can be proved by use of the following device of KAKEYA [5].

Consider the linear combinations

$$L_{\alpha,\beta}(z) = t_\alpha(z) + \lambda_1^{(\alpha)} t_{\alpha-1}(z) + \cdots + \lambda_{\alpha-\beta}^{(\alpha)} t_\beta(z), \qquad \alpha > \beta,$$

of the polynomials $t_\nu(z)$ introduced above where the λ's denote complex constants. Let δ denote the length of the greater of the diagonals of the parallelogram with the vertices $0, 1, 1+\vartheta, \vartheta$ (i.e., $\delta = \max(|1+\vartheta|, |1-\vartheta|)$). Among the considered linear combinations for every $\alpha > \beta$ we can find one in which the absolute values of all the factors λ are less than δ and the coefficients of the powers $z^{\alpha-1}, \ldots, z^\beta$ are integral numbers from $P(\vartheta)$ of the form $g+h\vartheta$ (with integral rational g and h). Indeed, every complex number if not itself coinciding with any of the above mentioned integral numbers, must lie inside or on the boundary of a parallelogram with vertices

$$g+h\vartheta, \qquad g+1+h\vartheta, \qquad g+1+(h+1)\vartheta, \qquad g+(h+1)\vartheta.$$

Thus we have the decomposition

$$L_{\alpha,\beta}(z) = G_{\alpha,\beta}(z) + H_{\alpha,\beta}(z) \tag{4}$$

where the coefficients $g_\nu^{(\alpha,\beta)}$ of

$$G_{\alpha,\beta}(z) = z^\alpha + g_1^{(\alpha,\beta)} z^{\alpha-1} + \cdots + g_{\alpha-1}^{(\alpha,\beta)} z + g_\alpha^{(\alpha,\beta)}$$

are integers of $P(\vartheta)$, while

$$H_{\alpha,\beta}(z) = h_0^{(\alpha,\beta)} z^\beta + h_1^{(\alpha,\beta)} z^{\beta-1} + \cdots + h_\beta^{(\alpha,\beta)}$$

is a polynomial whose degree is at most β and whose coefficients $h_\nu^{(\alpha,\beta)}$ have absolute values less than δ.

By (3) (and on account of $\gamma \geq 1$, $\varrho < 1$) $L_{\alpha,\beta}(z)$ must satisfy on S the following inequality

$$|L_{\alpha,\beta}(z)| \leq \gamma\,\delta(\varrho^\alpha + \varrho^{\alpha-1} + \cdots + \varrho^\beta) < \gamma\,\delta\,\frac{\varrho^\beta}{1-\varrho}. \tag{5}$$

Now let us fix the integer β so that

$$\gamma\,\delta\,\frac{\varrho^\beta}{1-\varrho} \leq \frac{1}{3}. \tag{6}$$

In virtue of $|h_\nu^{(\alpha,\beta)}| < \delta$, $\nu = 0, 1, \ldots, \beta$, we can select from the set of the polynomials

$$H_{\alpha,\beta}(z), \qquad \alpha = \beta+1, \beta+2, \ldots,$$

an infinite sequence $H_{\alpha_1,\beta}(z), H_{\alpha_2,\beta}(z), \ldots, H_{\alpha_j,\beta}(z), \ldots$ with $\beta<\alpha_1<\alpha_2<\cdots<\alpha_j<\cdots$, which converges uniformly in S as $j\to\infty$. Thus there is a pair of numbers $\alpha_\sigma, \alpha_\tau$ such that $\alpha_\sigma<\alpha_\tau$ and such that in S

$$|H_{\alpha_\sigma,\beta}(z) - H_{\alpha_\tau,\beta}(z)| \leq \tfrac{1}{3} \tag{7}$$

holds. Hence, by (4), (5) and (6),

$$|G_{\alpha_\tau,\beta}(z) - G_{\alpha_\sigma,\beta}(z)| \leq |L_{\alpha_\tau,\beta}(z)| + |L_{\alpha_\sigma,\beta}(z)| + |H_{\alpha_\tau,\beta}(z) - H_{\alpha_\sigma,\beta}(z)| < \\ < \tfrac{1}{3} + \tfrac{1}{3} + \tfrac{1}{3} = 1.$$

Consequently,

$$G_{\alpha_\tau,\beta}(z) - G_{\alpha_\sigma,\beta}(z) = z^{\alpha_\tau} + \ldots$$

possesses all the required properties of $c(z)$.

Thus the proof of Theorem B is complete.

§ 3. Theorems on Lemniscates.

1. First we shall deal with an almost obvious proposition:

C. *Let the bounded and closed set S contain a lemniscate*

$$|c(z)| = 1 \tag{8}$$

where the coefficients of the polynomial

$$c(z) = z^r + c_1 z^{r-1} + \cdots + c_{r-1} z + c_r \tag{9}$$

are rational integral numbers [or algebraic integral numbers from a quadratic-imaginary field $P(\vartheta)$]. Then there is an infinite number of equations of the class K [or of the class $K(\vartheta)$] whose roots all lie in S [namely, on the lemniscate-line (8)].

For the proof we note that, for $k=1, 2, \ldots$, all the equations

$$(c(z))^k - 1 = 0$$

have coefficients as $c(z)$ itself; the same is true for the equations

$$d_k(z) = \frac{(c(z))^k - 1}{c(z) - 1} = 0$$

where $k \geq 2$. All these equations have roots which lie on the lemniscate-line (8) and, consequently, also in S. Furthermore, $d_{k_1}(z)$ and $d_{k_2}(z)$ are certainly without common divisors whenever k_1 and k_2 denote different primes.

We denote by $\delta_2(z), \delta_3(z), \delta_5(z), \delta_7(z), \delta_{11}(z), \ldots$ factors of the polynomials $d_2(z), d_3(z), d_5(z), d_7(z), d_{11}(z), \ldots$, belonging to the class K [or to $K(\vartheta)$]. These $\delta(z)$'s yield infinitely many algebraic equations with all the properties required.

2. A result which lies much deeper than C and which forms the most relevant basis of our following considerations is

D. *Let*

$$b(z) = z^k + b_1 z^{k-1} + \cdots + b_k$$

be a polynomial with real (resp. arbitrary complex) coefficients and let R be any real number >1. Then there is a lemniscate (8): $|c(z)|=1$, *which is enclosed by the lemniscate*

$$|b(z)| = R. \tag{10}$$

Here the coefficients of $c(z)$ satisfy the conditions formulated in C.

We first prove the existence of a polynomial $\beta(z)=z^k+\beta_1 z^{k-1}+\cdots+\beta_k$ with coefficients of the fields P and $P(\vartheta)$, respectively, such that for suitably chosen R_1 and R_2 satisfying $1<R_1<R_2<R$, the lemniscate $|\beta(z)|=R_2$ encloses $|b(z)|=R_1$ but is enclosed by (10)[3].

Let $1<R_1<R$. Put

$$M = \max_{|b(z)|=1} (1+|z|+\cdots+|z|^{k-1}). \tag{11}$$

Choose $\varepsilon>0$ so that

$$\varepsilon M<1,\ \frac{1+\varepsilon M}{1-\varepsilon M} R_1<R, \tag{12}$$

and put $R_2=(1+\varepsilon M)R_1$. Clearly $R_1<R_2<R$. If the b's are real, choose rational $\beta_1,\ldots,\beta_k$ with $|\beta_j-b_j|<\varepsilon$. If the b's are complex, choose β_j in $P(\vartheta)$ satisfying the same inequalities.

Put

$$\beta(z) = z^k+\beta_1 z^{k-1}+\cdots+\beta_k. \tag{13}$$

Notice that, by (11),

$$\frac{1+|z|+\cdots+|z|^{k-1}}{|b(z)|} \leq M \tag{14}$$

for $|b(z)|=1$, and the fraction vanishes at ∞. Hence the inequality holds whenever $|b(z)|\geq 1$[4]. Thus, for $|b(z)|\geq 1$, we have

$$|\beta(z)-b(z)|<\varepsilon(|z|^{k-1}+\cdots+1)\leq \varepsilon M\,|b(z)|, \tag{15}$$

and therefore in particular

$$(1-\varepsilon M)\,|b(z)|<|\beta(z)|<(1+\varepsilon M)\,|b(z)|. \tag{16}$$

Consequently $|\beta(z)|<R_2$ for $|b(z)|=R_1$, and hence for $|b(z)|\leq R_1$. Hence if $|\beta(z)|=R_2$, then $|b(z)|>R_1$. On the other hand, under the same condition, in view of (16),

$$|b(z)|<\frac{1}{1-\varepsilon M}|\beta(z)| = \frac{1}{1-\varepsilon M} R_2<R. \tag{17}$$

Thus the lemniscate $|\beta(z)|=R_2$ lies in the region $R_1<|b(z)|<R$, enclosing $|b(z)|=R_1$ and being enclosed by $|b(z)|=R$ as required.

---

[3]) This proof of the existence of $\beta(z)$ (terminating at the end of 2.) which replaces a more complicated argument, is due to Professor R. M. Robinson to whom we owe also other improvements of the manuscript.

[4]) Cf. G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, Berlin, 1925, vol. 1, p. 142, problem 299.

3. We proceed now to the proof of Theorem D. The lemniscate $|\beta(z)| = R_2$ has a "radius" $R_2^{1/k}$ which exceeds 1. The polynomial (13) can be written as follows:

$$\beta(z) = z^k + \frac{\gamma_1 z^{k-1} + \gamma_2 z^{k-2} + \cdots + \gamma_k}{n}$$

where n denotes a natural number and the γ's are rational integers [resp. algebraic integers of $P(\vartheta)$]. Let first $n = 1$; then, because of the properties of $\beta(z)$ explained above,

$$|\beta(z)| = 1$$

will furnish the required lemniscate. In order to prove the existence of (8) also in the case $n > 1$, we choose a natural number μ (of which we shall dispose later) and we define $\sigma = \mu k$, $\nu = \sigma!\, n^\sigma$; obviously $\nu > \sigma \geq \mu$. Then we form

$$\begin{aligned}(\beta(z))^\nu = z^{k\nu} &+ \frac{\nu}{n} z^{k(\nu-1)}(\gamma_1 z^{k-1} + \cdots + \gamma_k) + \cdots \\ &+ \frac{\nu}{\sigma!\, n^\sigma}(\nu-1)\ldots(\nu-\sigma+1)\, z^{k(\nu-\sigma)}(\gamma_1 z^{k-1} + \cdots + \gamma_k)^\sigma + \\ &+ \binom{\nu}{\sigma+1} z^{k(\nu-\sigma-1)}\left(\frac{\gamma_1 z^{k-1} + \cdots + \gamma_k}{n}\right)^{\sigma+1} + \cdots.\end{aligned}$$

Here the coefficients of $z^{k\nu-1}, z^{k\nu-2}, \ldots, z^{k\nu-\sigma}$ are rational integers [resp. algebraic integers of $P(\vartheta)$]. Thus we can determine the real (resp. complex) numbers

$$\lambda_1^{(1)}, \ldots, \lambda_k^{(1)};\ \lambda_1^{(2)}, \ldots, \lambda_k^{(2)};\ \ldots;\ \lambda_1^{(\nu-\mu)}, \ldots, \lambda_k^{(\nu-\mu)}$$

such that in

$$\begin{aligned}\Gamma_\mu(z) = (\beta(z))^\nu &+ \lambda_1^{(1)} z^{k-1}(\beta(z))^{\nu-\mu-1} + \lambda_2^{(1)} z^{k-2}(\beta(z))^{\nu-\mu-1} + \cdots + \lambda_k^{(1)}(\beta(z))^{\nu-\mu-1} \\ &+ \lambda_1^{(2)} z^{k-1}(\beta(z))^{\nu-\mu-2} + \lambda_2^{(2)} z^{k-2}(\beta(z))^{\nu-\mu-2} + \cdots + \lambda_k^{(2)}(\beta(z))^{\nu-\mu-2} \\ &+ \cdots + \\ &+ \lambda_1^{(\nu-\mu)} z^{k-1} + \lambda_2^{(\nu-\mu)} z^{k-2} + \cdots + \lambda_k^{(\nu-\mu)} \\ = (\beta(z))^\nu &+ (\beta(z))^{\nu-\mu-1}[\lambda_1^{(1)} z^{k-1} + \lambda_2^{(1)} z^{k-2} + \cdots + \lambda_k^{(1)}] \\ &+ (\beta(z))^{\nu-\mu-2}[\lambda_1^{(2)} z^{k-1} + \lambda_2^{(2)} z^{k-2} + \cdots + \lambda_k^{(2)}] \\ &+ \cdots + [\lambda_1^{(\nu-\mu)} z^{k-1} + \lambda_2^{(\nu-\mu)} z^{k-2} + \cdots + \lambda_k^{(\nu-\mu)}] \\ = (\beta(z))^\nu &+ \Delta_\mu(z) = z^{k\nu} + \delta_1 z^{k\nu-1} + \cdots\end{aligned}$$

all coefficients are rational integers [resp. algebraic integers $P(\vartheta)$ of the form $g + h\vartheta$ with rational integers g and h]; also all λ's are in the absolute value < 1 (resp. $< 1 + |\vartheta|$) (see § 2.2).

For $|\beta(z)| = R_2$ it holds, since $R_2 > 1$,

$$\begin{aligned}\left|\frac{\Delta_\mu(z)}{(\beta(z))^\nu}\right| &< c \max_{|\beta(z)|=R_2} (1 + |z| + \cdots + |z|^{k-1})\left(\frac{1}{R_2^{\mu+1}} + \frac{1}{R_2^{\mu+2}} + \cdots + \frac{1}{R_2^\nu}\right) \\ &< c\,\frac{N}{R_2^\mu}\,\frac{1}{R_2 - 1}\end{aligned}$$

with $c=1$ (resp. $c=1+|\vartheta|$) and

$$N=\max_{|\beta(z)|=R_2}(1+|z|+\cdots+|z|^{k-1});$$

consequently, for $\mu\geq\mu_0$, we have

$$\left|\frac{\Delta_\mu(z)}{(\beta(z))^\nu}\right|<\frac{1}{2}.$$

Now, according to Rouché's theorem, all the zeros of $\Gamma_\mu(z)$ must lie in the interior of $|\beta(z)|=R_2$ as soon as $\mu\geq\mu_0$.

On the other hand, it holds, for $\mu\geq\mu_0$, everywhere on the same curve,

$$|\Gamma_\mu(z)|\geq|\beta(z)|^\nu-|\Delta_\mu(z)|=R_2^\nu\left(1-\frac{|\Delta_\mu(z)|}{R_2^\nu}\right)>\frac{R_2^\nu}{2}>\frac{R_2^\mu}{2}.$$

Thus, if $\mu_1\geq\mu_0$ is large enough, we have, for $\mu\geq\mu_1$, the inequality

$$|\Gamma_\mu(z)|>1; \tag{18}$$

hence the lemniscate

$$|\Gamma_\mu(z)|=1 \tag{19}$$

will be enclosed by $|\beta(z)|=R_2$, [and, therefore, *a fortiori*, by (10)]. Indeed, since the foci of (19) [*i.e.*, the zeros of $\Gamma_\mu(z)$] lie all, as already established, in the interior of $|\beta(z)|=R_2$, we can conclude from the minimum-principle of the theory of functions that everywhere in the exterior of $|\beta(z)|=R_2$ the inequality (18) must be valid. Thus the existence-proof of (8) is complete.

4. A combination of Theorems C and D yields

E. *If a bounded point-set of the complex z-plane contains a lemniscate-domain*

$$|z^k+b_1z^{k-1}+\cdots+b_k|\leq R \tag{20}$$

whose radius ($=R^{1/k}=$ *the transfinite diameter of this domain*) *exceeds unity, then there is an infinite number of algebraic equations of the class* $K(\vartheta)$ *whose roots all lie in* S.

Remark. According to C and D there is even an infinite number of equations of the characterized kind whose roots all lie on a lemniscate (8), which is enclosed by (20), where (9) has integral coefficients out of $P(\vartheta)$ (resp. rational integral coefficients, in the case that S is symmetrical with respect to the real axis). Thus we have already obtained a large class of point-sets S with an infinite number of equations of the required kind. Our further task is to replace the existence of the lemniscate-domain (20) by other conditions, which are easier to manage.

§ 4. Passage to more general sets.

1. The range of Theorem E and of the Remark formulated above can be recognized from the following proposition:

F. *Let S be a bounded and closed point-set of the z-plane, whose complementary set $\overline{S}$ is a connected open set. Denote by $S(\varrho)$ the set of points belonging to the circles with the radius ϱ described about the points of the set S.*

Let the transfinite diameter of S be not less than 1. Then however small $\varrho > 0$ may be, there is an infinite number of equations of the class $K(\vartheta)$ whose roots all lie in $S(\varrho)$. In the case where S is symmetrical with respect to the real axis of the z-plane, there is even an infinite number of equations of the class K whose roots all lie in $S(\varrho)$.

Remark to F. Under the above conditions regarding S, there is even an infinite number of equations of the characterized kind whose roots all lie on a lemniscate-curve (8) contained in $S(\varrho)$, where the coefficients of (9) are integral numbers of $P(\vartheta)$ (resp. rational integral numbers if S lies symmetrically to the real axis).

2. The proof of Theorem F and of the Remark can be based on the preceding, combined with an extension of Hilbert's classical theorem on the approximation of arbitrary analytical curves by lemniscates [4]. This extension, due to FEKETE [3, p. 638] (in a more restricted form than stated below), is as follows:

G. *Let S be an arbitrary set with the properties required at the beginning of F. Let $n \geq 2$; we denote by $\zeta_1^{(n)}, \ldots, \zeta_n^{(n)}$ a system of points on S for which $\Pi |z_i - z_k|$, $1 \leq i < k \leq n$, attains its maximum when $z_1, \ldots, z_n$ vary independently on S. Let μ_n denote the maximum of the absolute value of*

$$\tau_n(z) = (z - \zeta_1^{(n)}) \ldots (z - \zeta_n^{(n)})$$

in S. Then the lemniscate-domain, defined by

$$|\tau_n(z)| \leq \mu_n \tag{21}$$

is contained in $S(\varrho)$ provided $n \geq n_0 = n_0(\varrho)$. Here ϱ and $S(\varrho)$ have the same meaning as in F.

If S is symmetrical with respect to the real axis, then not only (21) but also the lemniscate-domain defined by

$$|\omega_n(z)| \leq \nu_n \tag{22}$$

is contained in $S(\varrho)$ whenever $n \geq n_1 = n_1(\varrho)$. Here $\omega_n(z)$ and ν_n are defined by

$$\omega_n(z) = \prod_{k=1}^{n} (z - \zeta_k^{(n)}) \cdot \prod_{k=1}^{n} (z - \overline{\zeta_k^{(n)}}) = \tau_n(z)\, \overline{\tau}_n(z),$$

$$\nu_n = \max_{z \text{ in } S} |\omega_n(z)|.$$

We take this Theorem G for granted and prove F as follows: If the set S of F has a transfinite diameter > 1, then, for $n \geq n_0(\varrho)$, (21) [resp., when S is symmetrical, for $n \geq n_1(\varrho)$, (22)] can undertake the role of the lemniscate-domain (20) of Theorem E, provided that the role of the point-set of E is passed on to $S(\varrho)$. In fact, the "radius" of (21) [$= \mu_n^{1/n} =$ the transfinite diameter of (21)] cannot, for any $n \geq 2$, be less than the transfinite diameter of the set S covered by (21), *i.e.*, it too must exceed unity for every $n \geq 2$ [resp., the radius of (22) must be greater than 1]; thus, for $n \geq n_0(\varrho)$ [resp.

$n \geq n_1(\varrho)$] (21) [resp., when S and, therefore, also $S(\varrho)$ are symmetrical, (22)] possesses all the properties required for (20).

In the case where the transfinite diameter of S is equal to 1, we can assert again, in view of the monotonity of the transfinite diameter as a set-function, the inequalities

$$\mu_n^{1/n} \geq 1, \quad \nu_n^{1/n} \geq 1$$

for every $n \geq 2$. Hence it follows, again with the aid of G, that for every $n \geq n_0(\varrho)$ [resp. $n \geq n_1(\varrho)$] and for sufficiently small non-negative $\varepsilon_n = \varepsilon_n(\varrho)$ [resp. $\eta_n = \eta_n(\varrho)$], the lemniscate-domain

$$|\tau_n(z)| \leq \mu_n + \varepsilon_n \tag{23}$$

[resp.

$$|\omega_n(z)| \leq \nu_n + \eta_n] \tag{24}$$

possesses all the properties required for (20) in Theorem E. Indeed, it has a radius greater than 1 and is completely contained in $S(\varrho)$. In fact, (21) [resp. (22)] is contained, under the above supposition, by virtue of G, for sufficiently large n, in $S(\varrho/2)$ and, consequently (23) [resp. (24)] can, by reason of continuity, have no point in the exterior of $S(\varrho)$ whenever $\varepsilon_n > 0$ (resp. $\eta_n > 0$) are sufficiently small. [Moreover $\omega_n(z)$ has real coefficients.]

Thus, by aid of the auxiliary proposition G, we have proved F and further, with regard to the remark to E, also the remark to F.

3. As to the auxiliary proposition G, we can prove it as follows. Let us assume that for a value of $\varrho > 0$ an infinite sequence

$$n_1 < n_2 < \cdots < n_k < \cdots \tag{25}$$

of natural numbers exists such that certain points $\zeta_{n_1}, \zeta_{n_2}, \ldots, \zeta_{n_k}, \ldots$ of the corresponding domains (21) lie outside of the set $S(\varrho)$ (consequently they lie in $\overline{S}$). Then, for a suitable subsequence

$$n_1', n_2', \ldots, n_k', \ldots \tag{26}$$

of (25), the corresponding ζ's would converge to a point ζ_0, situated in the finite portion of the z-plane (and belonging to $\overline{S}$) whose distance from S would be not less than ϱ.

From the existence of such a point ζ_0, however, a contradiction follows: Indeed, let C be a simple closed curve including the set S and excluding the circle $|z - \zeta_0| \leq \varrho/2$ (the existence of such a curve follows from the assumptions regarding S and from the definition of ζ_0). Let further $\delta = \max|z - t|$ and $\delta' = \min|z - t|$ where z ranges over the interior of C and t over the circle $|t - \zeta_0| \leq \varrho/2$, $\varrho > 0$. Then, according to the well-known theorem of RUNGE there is a polynomial $q(z; t)$ in z (with coefficients depending on t), such that

$$\left| \frac{1}{z - t} - q(z, t) \right| \leq \frac{1}{2\delta}$$

and, consequently,

$$|1 - (z - t)\, q(z, t)| \leq \tfrac{1}{2}$$

for all z inside or on C and for all t subject to $|t-\zeta_0| \leq \varrho/2$. Further, we define the polynomial (cf. SZEGÖ [7])

$$P_n(z;t) = \frac{1-[1-(z-t)\,q(z,t)]^\nu}{t-z}, \qquad n = 1, 2, 3, \ldots,$$

in z, where ν denotes the greatest integer in n/h and $h-1$ the degree of $q(z,t)$ in z. We have evidently the following properties:

(1) the degree of $P_n(z;t)$ in z is $\leq n-1$;

(2) for all z in S, for all t of the circle $|t-\zeta_0| \leq \varrho/2$ and for all $n \geq 2$, the inequality

$$\left|\frac{1}{t-z} - P_n(z;t)\right| \leq c\,\Theta^n \tag{27}$$

holds where Θ denotes a certain number between 0 and 1 and c a suitable positive constant. [One can set $c = 2/\delta'$, $\Theta = 2^{-1/n}$.]

Now we form the LAGRANGE interpolation polynomial $F_n(z,t)$ associated with the function $(t-z)^{-1} - P_n(z;t) = f_n(z;t)$ and with the points of interpolation $z = \zeta_1^{(n)}, \ldots, \zeta_n^{(n)}$. Because of the supposed extremal properties of these points and by (27) (valid for the values z, t, n specified there) we have

$$|F_n(z;t)| \leq n\,c\,\Theta^n \tag{28}$$

Further, the representation

$$F_n(z;t) = \frac{1}{t-z} - P_n(z;t) - \frac{\tau_n(z)}{(t-z)\,\tau_n(t)} \tag{29}$$

holds, since the right side is a polynomial in z of at most of the degree $n-1$ which coincides for $z = \zeta_\nu^{(n)}$ $(1 \leq \nu \leq n)$ with the above function $f_n(z,t)$ and, consequently with the left side of (29). From (27), (28) and (29) we conclude the relation

$$\left|\frac{\tau_n(z)}{\tau_n(t)}\right| \leq |t-z|\,(n+1)\,c\,\Theta^n \tag{30}$$

for all values of n, z and t defined above.

The announced contradiction becomes evident from (30) when one chooses for n sufficiently large numbers of the sequence (26) and for z points of the set S such that $|\tau_n(z)| = \mu_n$ holds; finally we set $t = \zeta_n$. In fact, under these conditions we have $|\tau_n(\zeta_n)| \leq \mu_n$, so that the left hand side of (30) is not less than unity, while the right hand side becomes arbitrarily small.

Thus we have proved the first assertion of proposition G.

The second assertion is also easily verified. Assuming that for a value of ϱ and for a sequence (25) the points $\zeta_{n_1}, \ldots, \zeta_{n_k}, \ldots$ of the lemniscate regions (22) would lie outside $S(\varrho)$, the ζ's would converge, for a suitable subsequence (26) of (25), to a (finite) point ζ_0 of the region $\overline{S}$ whose distance from S would not be less than ϱ. From the existence of such a point we can again derive a contradiction whenever S lies symmetrically to the real axis of the z-plane. In fact, one obtains, as above, with the same values of n, z and t, the inequality (30). Also, by the symmetry of S with respect to the real axis, together with

$\zeta_1^{(n)}, \ldots, \zeta_n^{(n)}$ also their conjugate values $\overline{\zeta_1^{(n)}}, \ldots, \overline{\zeta_n^{(n)}}$ make $\Pi|z_i - z_k|$ to a maximum; hence the polynomials $\bar{\tau}_n(z) = \prod_{k=1}^{n} (z - \overline{\zeta_k^{(n)}})$ must satisfy the inequalities

$$\left|\frac{\bar{\tau}_n(z)}{\bar{\tau}_n(t)}\right| \leq |t - z|\,(n+1)\,c\,\Theta^n. \tag{31}$$

But, from (30) and (31) it follows (for the above values of z, t and n):

$$\left|\frac{\omega_n(z)}{\omega_n(t)}\right| \leq |t - z|^2\,(n+1)^2\,c^2\,\Theta^{2n};$$

choosing z from S with $|\omega_n(z)| = \nu_n$, $t = \zeta_n$ and $n = n_k'$ (for sufficiently large k), this inequality leads to a contradiction.

5. It is not without interest to observe that, while the second assertion of the remark to F does not yet follow from the first one, the second part of F itself can be deduced from the first one. Indeed, let S be symmetrical with respect to the real axis; one can always specify an infinite sequence $a_1(z), \ldots, a_\nu(z), \ldots$ of the polynomials of the class $K(\vartheta)$ so that all their zeros are in $S(\varrho)$; one can even assume that for the said sequence not only $a_\nu(z) \neq a_\mu(z)$, but also $a_\nu(z) \neq \bar{a}_\mu(z)$, whenever $\nu \neq \mu$. Under this assumption either the sequence in question or the sequence $a_1(z)\,\bar{a}_1(z), \ldots, a_\mu(z)\,\bar{a}_\mu(z), \ldots$ contains an infinite subsequence of polynomials belonging to K; in the latter case the zeros lie all in $S(\varrho)$ provided S is symmetrical to the real axis.

§ 5. Extensions.

1. The following theorem provides an essential extension of our result formulated in the former paragraph.

H. *The assertions of Theorem F (and of the subsequent remark) remain valid when the requirement concerning the complementary set $\bar{S}$ of S is omitted.*

The proof of this theorem which forms a true counterpart to the Theorems A and B depends, besides Theorem F (and the remark to F), on two auxiliary theorems. The first of them (which plays here a similar role as G played in the proof of F) reads as follows:

I. *Let S be a bounded and closed set in the z-plane; the quantities n, $\tau_n(z)$, μ_n, $\omega_n(z)$, ν_n, ϱ, $S(\varrho)$ shall have the same meaning as in G. Then the lemniscate*

$$|\tau_n(z)| = \mu_n \tag{32}$$

[resp., when S is symmetrical with respect to the real axis, the lemniscate

$$|\omega_n(z)| = \nu_n] \tag{33}$$

runs completely in $S(\varrho)$ provided $n \geq n_0(\varrho)$ [resp., $n \geq n_1 = n_1(\varrho)$].

The second auxiliary theorem yields a continuity property of the transfinite diameter which is interesting in itself.

J. *Let S be a set as in I, with the transfinite diameter $d = d(S)$, and let z_0 be a point of S. Let S_1 denote that subset of S whose points z satisfy the*

inequality $|z-z_0| \geq \delta$. *If* $\varepsilon > 0$ *is an arbitrary small quantity,* S_1 *has a transfinite diameter* $d(S_1) > d(S) - \varepsilon$ *whenever* $\delta = \delta(\varepsilon)$ *is chosen sufficiently small.*

2. We first establish, relying upon I and J, the proof of H. Let the transfinite diameter of S be not less than 1 (resp., let, in addition, S be symmetrical to the real axis). By virtue of the auxiliary Theorem I, the lemniscate

$$|\tau_n(z)| = \mu_n + \varepsilon_n \tag{34}$$

[resp.

$$|\omega_n(z)| = \nu_n + \eta_n] \tag{35}$$

will be contained, for $n \geq n_0(\varrho/2)$ [resp. for $n \geq n_1(\varrho/2)$] and, for a suitably chosen non-negative ε_n (resp. η_n), in $S(3\varrho/4)$ and thereby will have a radius exceeding unity (compare the results of § 4.2); this lemniscate will be either a single simple closed curve K or it will consist of ν separated curves K_μ ($1 \leq \mu \leq \nu$; $2 \leq \nu \leq n$ resp. $\leq 2n$). (If S is *symmetrical* to the real axis, the same will be true of K or the system of curves K_μ.)

Let us fix now a point z_0 on K (this being a real point when S is symmetrical), resp. a point z_μ on each curve K_μ (when S is symmetrical, these points are to be chosen either real, or as pairs of points lying symmetrically to the real axis); the positive number δ can be chosen so small that the set Γ which results from K resp. K_μ through omission of all the points z satisfying $|z-z_0| < \delta$ resp. $|z-z_\mu| < \delta$, $1 \leq \mu \leq \nu$, yields a subset of (34) [resp. of (35)] with a transfinite diameter still greater than 1. Thereby Γ will have, as easily seen, an open and connected set as complementary set (when S is symmetrical, Γ will also be, according to our assumption, symmetrical to the real axis).

Applying now F (and the remark to F) to the set Γ, we obtain the existence of infinitely many equations with the property characterized there whose roots lie in $\Gamma(\varrho/4)$ and, *a fortiori*, in $S(\varrho)$ (and even on lemniscate-curves of the kind specified in the remark). Thus the proof of H is complete.

3. We can now be brief in proving I. For the part of the assertion concerning the lemniscate (32), we can proceed again as in the proof of G. As to the assertion about the lemniscate (33), we shall restrict ourselves to the remark that this assertion, as well as the corresponding one in G about the lemniscate-domain (22) bounded by (33), can be proved similarly as in § 4.4.

4. In order to prove J, we suppose $d(S) > 0$. Let $\delta > 0$ be so small that $2\delta < d(S)$ and $2\delta < 1$. We consider, for every $n \geq 2$, the points $\zeta_1^{(n)}, \ldots, \zeta_n^{(n)}$ of the set S for which the absolute value of $\Pi(z_i - z_k)$, $1 \leq i < k \leq n$, attains its maximum $d_n^{\binom{n}{2}} = [d_n(S)]^{\binom{n}{2}}$ when $z_1, \ldots, z_n$ vary independently in S. We divide these points in two classes, C_1 and C_2, according as they belong to the subset S_1 of S or not. Let $n_1 = n_1(n)$ be the number of the points in C_1, $n_2 = n_2(n)$ that of the points in C_2; $n_1 + n_2 = n$. Evidently

$$[d_n(S)]^{\binom{n}{2}} \leq [d_{n_1}(S_1)]^{\binom{n_1}{2}} (2\delta)^{\binom{n_2}{2}} [d_2(S)]^{n_1 n_2} \tag{36}$$

where $d_2(S)$ denotes the ordinary diameter (span) of S. In the case $0 \leq n_1 \leq 1$ the first factor on the right is to be replaced by 1.

We assert, first of all, that the sequence $n_1 = n_1(n)$, $n = 2, 3, \ldots$ is not bounded. In the contrary case one could deduce from (36), by passing to the limit $n \to \infty$, the inequality $d(S) \leq 2\delta$, which would contradict the above hypothesis about δ. Let now $n \to \infty$ over an arbitrary infinite sequence of natural numbers for which $n_1(n) \to \infty$ and let the limits

$$\lim \frac{n_1}{n} = \alpha, \qquad \lim \frac{n_2}{n} = \beta, \qquad 0 \leq \alpha = 1 - \beta \leq 1, \tag{37}$$

exist. Then, we conclude from (36) that $\alpha > 0$ and $d(S_1) > 0$ satisfying

$$d(S) \leq [d(S_1)]^{\alpha^2} (2\delta)^{\beta^2} [d_2(S)]^{2\alpha\beta}. \tag{38}$$

[In deed $\alpha = 0$ would imply, by (36), $d(S) \leq 2\delta$, against hypothesis.] If $\alpha = 1$, $\beta = 0$, then (38) entails $d(S) \leq d(S_1)$, i.e., $d(S) = d(S_1)$, and thus our assertion. Let us therefore consider the case $0 < \alpha < 1$ and, consequently, $1 > \beta > 0$. Then (38) yields

$$d(S) \leq [d(S)]^{\alpha^2} (2\delta)^{\beta^2} [d_2(S)]^{2\alpha\beta},$$

i.e.,

$$[d(S)]^{1+\alpha} \leq (2\delta)^{\beta} [d_2(S)]^{2\alpha}.$$

Hence

$$\beta \log \frac{1}{2\delta} \leq 2\alpha \log d_2(S) - (1+\alpha) \log d(S) \leq 2 |\log d_2(S)| + 2 |\log d(S)|. \tag{39}$$

This inequality shows that β can be made arbitrarily small if δ is small enough. The same follows by (39) for $\beta^2 \log \frac{1}{2\delta}$, so that $(2\delta)^{\beta^2}$ comes arbitrarily close to 1. The same holds for the third factor on the right of (38) while the first one approaches arbitrarily close to $d(S_1) > 0$. Thus our assertion is proved.

5. We derive now an immediate corollary of H providing a simple but very useful sufficient condition for the existence discussed in F and H. This corollary is as follows:

K. *Let a bounded and closed set S contain a closed subset S_1 whose points are all inner points of S, and whose transfinite diameter is at least 1. There are infinitely many equations of the class $K(\vartheta)$ whose roots lie in S.*

If S_1 is symmetrical with respect to the real axis of the z-plane, then there are infinitely many equations of the class K whose roots lie in S.

Remark. There exist infinitely many equations whose coefficients satisfy the conditions in the first (resp., when S_1 is symmetrical to the real axis, in the second) part of K and whose roots lie on a lemniscate of the form (8) contained in $S_1(\varrho)$, however small the positive quantity ϱ may be. The polynomial (9) possesses in this case integral coefficients only belonging to $P(\vartheta)$ (resp., when S_1 is symmetrical, to the rational number field).

In order to prove K (and the remark) one shall take into consideration the fact that, according to the assumption, with every point P of S_1 also a circle around it belongs to S. If, however, r_P denotes the radius of the largest

possible circle of this kind and ϱ_0 the minimum of r_P while P ranges over all points of S_1, then ϱ_0 must turn out positive. Thus, together with S_1, also $S_2 = S_1(\varrho)$ must be wholly contained in S, provided $0 < \varrho \leq \varrho_0$.

Thus H and the remark to H yield (when applied to S_1), Theorem K, and the remark to K whenever $0 < \varrho \leq \varrho_0$.

6. In conclusion an interesting corollary of K can be stated as Theorem

L. *Let $\Delta \ni \infty$ be a simply or multiply connected domain in the z-plane and Γ its boundary with $d(\Gamma) > 1$. Suppose that the complement E of $\Delta + \Gamma$ consists of a finite*[5]) *set of disjoint simply connected domains, say D_ϱ, $1 \leq \varrho \leq r$. Then there are infinitely many equation of the class $K(\vartheta)$ whose roots lie in E and there are even infinitely many equations of this class whose roots lie all on a lemniscate* (8), *the polynomial* (9) *possessing integral coefficients which belong to $P(\vartheta)$.*

For the proof we combine Theorem K with the following proposition due to FEKETE[6]):

Let Δ, Γ, and E satisfy the conditions of Theorem L. Given an $\varepsilon > 0$, there is a lemniscate Λ having r branches, say Λ_ϱ, $1 \leq \varrho \leq r$, so that Λ_ϱ is in D_ϱ and moreover such that

$$1 \leq d(\Gamma)/d(\Lambda) \leq 1 + \varepsilon.$$

Bibliography.

[1] BACHMANN, P.: Allgemeine Arithmetik der Zahlenkörper. Leipzig 1905. — [2] FEKETE, M.: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen. mit ganzzahligen Koeffizienten. Math. Z. **17**, 228—249 (1923). — [3] FEKETE, M.: Über den transfiniten Durchmesser ebener Punktmengen. (Dritte Mitt.) Math. Z. **37**, 635—646 (1933). — [4] HILBERT, D.: Über die Entwickelung einer beliebigen analytischen Funktion einer Variablen in eine unendliche nach ganzen rationalen Funktionen fortschreitenden Reihe. Nachr. Ges. Wiss. Göttingen, math.-phys. Kl. **1897**, 63—70. — [5] OKADA, Y.: On approximate polynomials with integral coefficients only. Tôhoku Math. J. **23**, 26—35 (1924). — [6] SCHUR, I.: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. Math. Z. **1**, 377—402 (1918). — [7] SZEGÖ, G.: Tschebyscheffsche Polynome und nichtfortsetzbare Potenzreihen. Math. Ann. **87**, 90—111 (1922).

[5]) This supposition of finiteness is essential to the subsequent conclusion. Eventually, E might fail to harbour any unit lemniscate (8), the polynomial (9) possessing integral coefficients of a field $P(\vartheta)$, or even to enclose all roots of a single equation (2) of the class $K(\vartheta)$ when it consists of infinetely many (disjoint) simply connected domains D_ϱ, $\varrho = 1, 2, 3, \ldots$.

[6]) On the semi-continuity of the transfinite diameter. Bull. Research Council of Israel **3** (1954), p. 336.

The Hebrew University, Jerusalem (Israel).
Stanford University, Stanford, Cal. (USA.).

(Eingegangen am 7. März 1955.)

BERGMAN, S.
Math. Zeitschr. Bd. 63, S. 173—194 (1955)

Bounds for Analytic Functions in Domains with a Distinguished Boundary Surface *).

To the memory of I. SCHUR.

By

STEFAN BERGMAN.

I. Introduction.

When one investigates the connection between value distributions of an analytic function inside a given domain and on its boundary, a basic difference arises between the theory of functions of one complex variable and of functions of several complex variables. When considering such relations in the case of two complex variables, it is useful to study at first special domains, namely domains bounded by finitely many segments of analytic hypersurfaces. On the (three-dimensional) boundary $\mathfrak{m}^3$ of such a domain $\mathfrak{M}$ lies the (two-dimensional) distinguished boundary surface $\mathfrak{D}^2$ which to a certain extent plays a similar role as the boundary curve does in the case of a plane domain. (See II, § 1.) Generalizing the CAUCHY formula in [2, 4, 5, 19, 20] [1]) A. WEIL and the author gave formulas expressing the value of a regular function $f(z_1, z_2)$ inside $\mathfrak{M}$ in terms of the value of f on $\mathfrak{D}^2$. (See also SOMMER [17].) However, this formula is more complicated than in the case of one complex variable, since its kernel (i.e., the analog of $[2\pi(z-t)]^{-1}$) depends upon the domain. It is therefore of interest to derive other representations of functions $f(z_1, z_2)$ inside $\mathfrak{M}$ in terms of its values on the distinguished boundary surface. In the case of one variable SZEGÖ [18] introduced functions which are orthonormal when integrating along the boundary of the domain. In section II, § 3 of the present paper we consider functions which are orthonormal when integrating over the distinguished boundary surface. They yield a new representation of $f(z_1, z_2)$ in terms of the values on the distinguished boundary surface. Some applications of these representations to the theory of regular and meromorphic functions are discussed in II, § 5.

In II, § 6 we extend the method of the minimum integral to the case of functions for which the integral is finite when integrating over the distinguished boundary surface (for details see II, § 6). The method of the minimum integral is a general procedure to obtain distortion theorems, see [7], [8].

In section III we consider relations between the value distribution of the function f in $\mathfrak{M}$ and in the complementary part $\mathfrak{b}^3 = \mathfrak{m}^3 - \mathfrak{D}^2$ of the boundary. (Here we are particularly interested in relations which have a different character than those in the case of one variable.) In our case $\mathfrak{m}^3$ is a union

*) Work done under contract with the Office of Naval Research.

[1]) Numbers in brackets refer to the bibliography.

of (finitely many) segments $\mathfrak{i}_k^3$ of analytic hypersurfaces[2]). To every $\mathfrak{i}_k^3$ we can determine a subdomain $\mathfrak{E}$ of $\mathfrak{R}$ possessing the following properties. If a function f is bounded on a (conveniently chosen) line $\mathfrak{j}_k^1 < \mathfrak{i}_k^3$ and omits in $\mathfrak{i}_k^3$ two values, say 0 and 1, then there exists a function $U(z_1, z_2, \mathfrak{j}_k^1)$ depending only upon the domain $\mathfrak{R}$ and the line $\mathfrak{j}_k^1$ so that $|f(z_1, z_2)| \leq U(z_1, z_2, \mathfrak{j}_k^1)$, $(z_1, z_2) \in \mathfrak{E}$.

Each $\mathfrak{i}_k^3$ is a sum $\mathfrak{i}_k^3 = \mathfrak{S} \mathfrak{F}_k^3(\lambda_k)$ of segments of analytic surfaces. In III, § 3 we consider analogous question assuming that (instead of omitting two values) $f(z_1, z_2)$ assumes in every $\mathfrak{F}_k^3(\lambda_k)$ every value at most once.

Notation. The manifolds are denoted as a rule by German characters. The upper index indicates the dimension of the manifold. In the case of four-dimensional manifolds the index 4 is omitted. (Also, the index 0 is not indicated.)

The author wishes to express his thanks to Mr. M. MASCHLER for his helpful suggestions in connection with the present paper.

II. A representation of a function of two complex variables in terms of its value on the distinguished boundary surface.

§ 1. Introductory remarks. In the case of one complex variable, in [*1, 11*] function systems which are orthonormal when integrating over the domain were derived, while in [*18*] systems which are orthonormal when integrating along the boundary curve were introduced. When we pass to the case of functions of several complex variables, in the case of domains with a distinguished boundary surface, in addition to the two analogous possibilities (integration over the whole domain and over the boundary), we can introduce functions where the integration is carried out over the distinguished boundary surface.

DEFINITION. A (smallest) two-dimensional set $\mathfrak{D}^2$ (not necessarily connected) lying on the boundary $\mathfrak{m}^3$ of $\mathfrak{R}$, and possessing the property that every function f regular in the closed domain assumes the maximum of its absolute values, at least at one point of $\mathfrak{D}^2$, is called the distinguished boundary surface of $\mathfrak{R}$.

In particular, if the boundary $\mathfrak{m}^3$ is a sum of segments of analytic hypersurfaces, then the distinguished boundary surface is the sum the of the intersections of these segments. See also (2.6a).

In the case of domains which are bounded by two segments of analytic hypersurfaces, an integral formula where integration is taken over the distinguished boundary surface, has been derived in [*3*]. In the present paper the same problem for domains bounded by finitely many analytic hypersurfaces, i.e., with distinguished boundary surface with more complicated

[2]) In the present paper we consider only *one* segment of an analytic surface intersecting our domain along a line $\mathfrak{a}^1$ lying in $\mathfrak{i}_k^3$. Forming a (two parameter) family of such segments one can obtain the subdomains $\mathfrak{E}$. The question of choosing the subdomain $\mathfrak{E}$, to get "best" bounds $U(z_1, z_2, \mathfrak{j}_k^1)$ will be discussed in a later paper.

structure is considered. We shall assume that the domains under consideration (which will be denoted by $\mathfrak{R}$) are of a comparatively simple structure. We wish however to stress that these simplifications are made only to shorten the formulation of our hypotheses and that essentially the same results are valid in the case of general domains considered in [4, 5].

§ 2. Domains $\mathfrak{R}$ and their structure[2a]). We assume that a domain $\mathfrak{R}$ is bounded by finitely many (closed) segments $\overline{\mathfrak{i}}_k^3$, $k=1, 2, \ldots, n$, of analytic hypersurfaces. Every $\overline{\mathfrak{i}}_k^3 = \mathop{\mathsf{S}}_{\lambda_k \in \mathfrak{s}^1} \overline{\mathfrak{J}}_k^2(\lambda_k)$, where $\mathfrak{s}^1 = [0 \leq \lambda_k < 2\pi]$. $\overline{\mathfrak{J}}_k^2(\lambda_k)$ is a domain lying in the analytic surface

$$[\Phi_k(z_1, z_2, \lambda_k) = 0],$$

$\Phi_k(z_1, z_2, \lambda_k)$ being analytic functions of two complex variables z_1, z_2 which are regular in $\mathfrak{R}$ and continuously differentiable in λ_k (Hypothesis **1, a**). We assume further that $\overline{\mathfrak{J}}_k^2(\lambda_k)$ can be uniformized, i.e. to every $\Phi_k(z_1, z_2, \lambda_k)$ there exist two continuously differentiable functions $z_\nu = h_k^\nu(Z_k, \lambda_k)$, $\nu=1, 2$, defined in a domain $[Z_k \in \overline{\mathfrak{B}}_k^2(\lambda_k)]$ which map it $(1-1)$ onto $\overline{\mathfrak{J}}_k^2(\lambda_k)$ such that

$$\Phi_k[h_k^1(Z_k, \lambda_k),\ h_k^2(Z_k, \lambda_k),\ \lambda_k] = 0, \qquad Z_k \in \overline{\mathfrak{B}}_k^2(\lambda_k). \tag{2.1}$$

For every fixed λ_k the $h_k^\nu(Z_k, \lambda_k)$ are analytic functions of the complex variable Z_k, which are regular in the domain $\overline{\mathfrak{B}}_k^2(\lambda_k)$ of the Z_k plane. Every $\overline{\mathfrak{B}}_k^2(\lambda_k)$ is a simply connected (closed) domain which is bounded by a rectifiable curve (Hypothesis **1, b**). As we mentioned before, the boundary $\mathfrak{m}^3$ of $\mathfrak{R}$ consists of n segments of analytic hypersurfaces $\overline{\mathfrak{i}}_k^3$,

$$\mathfrak{m}^3 = \mathop{\mathsf{S}}_{k=1}^{n} \overline{\mathfrak{i}}_k^3. \tag{2.2}$$

We distinguish on $\mathfrak{m}^3$ two kinds of points. *I-points* i.e. points which belong to *one* $\overline{\mathfrak{i}}_k^3$ only and *K-points*, i.e., points which belong to the *intersection of two or more* $\overline{\mathfrak{i}}_k^3$'s. According to [4, 5], for every K-point (z_1, z_2) which lies on the intersection of $\overline{\mathfrak{i}}_k^3$ and $\overline{\mathfrak{i}}_s^3$ we have

$$z_m = h_k^m(Z_k, \lambda_k) = h_s^m(Z_s, \lambda_s), \qquad m = 1, 2. \tag{2.3}$$

We assume that

$$\frac{\partial[h_k^1(Z_k, \lambda_k) - h_s^1(Z_s, \lambda_s),\ h_k^2(Z_k, \lambda_k) - h_s^2(Z_s, \lambda_s)]}{\partial(\lambda_k, \lambda_s)} \neq 0 \tag{2.4}$$

holds (Hypothesis **2, d**).

By repeating the considerations of [4] one can show that all points of boundary curves $\mathfrak{i}_k^1(\lambda_k)$ of every $\overline{\mathfrak{J}}_k^2(\lambda_k)$ must be K-points. Since every point of $\mathfrak{i}_k^1(\lambda_k)$ must belong to an $\overline{\mathfrak{i}}_s^3$, $s \neq k$, we can write

$$\mathfrak{i}_k^1(\lambda_k) = \mathop{\mathsf{S}}_{\substack{s=1 \\ s \neq k}}^{n} \mathfrak{i}_{ks}^1(\lambda_k), \qquad \mathfrak{i}_{ks}^1(\lambda_k) = \mathfrak{i}_k^1(\lambda_k) \cap \overline{\mathfrak{i}}_s^3, \tag{2.5}$$

[2a]) The hypotheses in § 2 are formulated in a rather general way and thus appear complicated. Since these elaborate hypotheses are not essential for the understanding of the basic ideas of the paper, the reader may skim over them in the first reading.

and

$$\mathfrak{G}^2_{k\,s} = \sum_{\substack{\lambda_k \in s^1 \\ k \neq s}} \mathfrak{i}^1_{k\,s}(\lambda_k) = \sum_{\lambda_k \in s^1} \mathfrak{i}^1_{s\,k}(\lambda_s). \tag{2.6}$$

$$\mathfrak{D}^2 = \sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \mathfrak{G}^2_{k\,s} \tag{2.6a}$$

is the distinguished boundary surface of $\mathfrak{R}$.

Applying the theorem on implicit functions and that of HEINE-BOREL we can show that we can cover every $\mathfrak{G}^2_{k\,s}$ by denumerably many neighborhoods $\mathfrak{R}^2_{k\,s,\mu}$ (of the λ_k, λ_s-plane), $\mu = 1, 2, 3, \ldots$ such that in every $\mathfrak{R}^2_{k\,s,\mu}$ the corresponding part of $\mathfrak{G}^2_{k\,s}$ can be represented in the form

$$z_m = \varphi_m(\lambda_k, \lambda_s) \equiv h^m_k[Z^*_k(\lambda_k, \lambda_s), \lambda_k] \equiv h^m_s[Z^*_s(\lambda_s, \lambda_k), \lambda_s], \quad m = 1, 2, \quad k \neq s. \tag{2.7}$$

Here Z^*_k and Z^*_s are single valued and continuously differentiable functions of λ_k and λ_s in a conveniently chosen domain $\mathfrak{U}^2_{k\,s,\mu}$ of the λ_k, λ_s-plane. The domain $\sum_{\mu=1}^{\infty} \mathfrak{U}^2_{k\,s,\mu}$ is a possibly completely or partially multiplicovered RIEMANN's surface whose projection on the λ_k, λ_s-plane lies in

$$[0 \leq \lambda_k \leq 2\pi, \qquad 0 \leq \lambda_s \leq 2\pi].$$

We write

$$\sum_{\mu=1}^{\infty} \mathfrak{U}^2_{k\,s,\mu} = \mathfrak{B}^2_{k\,s} = \sum_{\lambda_k \in s^1} \mathfrak{u}^1_{k\,s}(\lambda_k) = \sum_{\lambda_s \in s^1} \mathfrak{u}^1_{s\,k}(\lambda_s). \tag{2.8}$$

Here $\mathfrak{u}^1_{k\,s}(\lambda_k)$ [the image of $\mathfrak{i}^1_{k\,s}(\lambda_k)$ in the λ_k, λ_s-plane], is completely or partially multiply covered set of $[0 \leq \lambda_s \leq 2\pi]$.

We assume that $\mathfrak{B}^2_{k\,s}$ is measurable and that $\left|\frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)}\right|$ for $(\lambda_k, \lambda_s) \in \mathfrak{B}^2_{k\,s}$ is bounded and the area of $\mathfrak{B}^2_{k\,s}$ is finite (Hypothesis **2, c***). (This assumption is stronger than the hypothesis **2, c** of [5]).

Further we assume that there exists a domain say $\mathfrak{A}^2$ in the z_2 plane with the following properties: For $t_2 \in \mathfrak{A}^2$, $\mathfrak{R} \cap (z_2 = t_2)$ is not empty and its boundary is $\mathfrak{m}^3 \cap (z_2 = t_2)$ (Hypothesis **3, a**). Every $\mathfrak{m}^3 \cap (z_2 = t_2)$, $t_2 \in \mathfrak{A}^2$, consists of finitely many I-arcs and finitely many K-points (Hypothesis **3, b**).

Let[3] $Z_k = T^{(\mu)}_k \equiv T^{(\mu)}_k(\lambda_k, t_2)$, $\mu = 1, 2, \ldots$, be the solution of the equation

$$h^2_k(Z_k, \lambda_k) = t_2, \qquad t_2 \in \mathfrak{A}^2, \tag{2.9}$$

and let

$$z_1 = \zeta_1(\lambda_k, t_2) \equiv h^1_k[T^{(\mu)}_k(\lambda_k, t_2), \lambda_k]. \tag{2.9a}$$

We assume that to the increasing λ_k corresponds the direction along the arc (2.9a) so that the domain $\mathfrak{R} \cap (z_2 = t_2)$ lies to the left (Hypothesis **3, c**).

[3] $T^{(\mu)}_k(\lambda_k, t_2)$ is in general a multivalued function, and μ indicates the branch. However, we shall omit in the future sometimes the index μ, e.g. we write ζ_1 instead of $\zeta^{(\mu)}_1$.

For $\frac{\partial \zeta_1(\lambda_k, t_2)}{\partial \lambda_k}$ the following representation holds:

$$\frac{\partial \zeta_1}{\partial \lambda_k} = \frac{\partial h_k^1}{\partial T_k^{(\mu)}} \frac{\partial T_k^{(\mu)}}{\partial \lambda_k} + \frac{\partial h_k^1}{\partial \lambda_k} = \frac{\partial (h_k^1, h_k^2)}{\partial (\lambda_k, T_k^{(\mu)})} : \frac{\partial h_k^2}{\partial T_k^{(\mu)}} \tag{2.10}$$

since from (2.9) we have:

$$\frac{\partial h_k^2}{\partial T_k^{(\mu)}} \frac{\partial T_k^{(\mu)}}{\partial \lambda_k} + \frac{\partial h_k^2}{\partial \lambda_k} = 0.$$

According to **3, b** for every t_2, $t_2 \in \mathfrak{A}^2$, there exist at most finitely many K-points which lie on $\mathfrak{G}_{k,s}^2 \cap (z_2 = t_2)$. For these points

$$h_k^2(Z_k, \lambda_k) = h_s^2(Z_s, \lambda_s) = t_2 \tag{2.11}$$

holds. We assume (Hypothesis **3, d**) that in these points:

$$\frac{\partial h_k^2(Z_k, \lambda_k)}{\partial Z_k} \neq 0, \quad \frac{\partial h_s^2(Z_s, \lambda_s)}{\partial Z_s} \neq 0. \tag{2.12}$$

Let further the images of *these* K-points in the λ_k, λ_s-plane be the points

$$M^p \equiv \{\lambda_k^{(p)}, \lambda_s^{(p)}\}, \qquad p = 1, 2, \ldots, l.$$

In a sufficiently small λ_k, λ_s-neighborhood we have for $\varphi_2(\lambda_k, \lambda_s)$ the representation [see (2.7)]

$$\varphi_2 = t_2 + A_1(\lambda_k, \lambda_s)(\lambda_k - \lambda_k^{(p)}) + A_2(\lambda_k, \lambda_s)(\lambda_s - \lambda_s^{(p)}) \tag{2.13}$$

where $A_m(\lambda_k, \lambda_s)$, $m = 1, 2$ are continuous functions defined in $\mathfrak{U}_{k,s,p}^2$. We assume (Hypothesis **3, e**) that there exist constants A^* and Θ^* such that

$$\min [|A_1(\lambda_k, \lambda_s)|, \quad |A_2(\lambda_k, \lambda_s)|] \geq A^* > 0, \tag{2.14}$$

$$\pi - \Theta^* \geq |\arc A_1(\lambda_k, \lambda_s) - \arc A_2(\lambda_k, \lambda_s)| \geq \Theta^* > 0, \quad (\lambda_k, \lambda_s) \in \mathfrak{U}_{k,s,p}^2. \tag{2.15}$$

Under the above hypotheses it is possible to show that every analytic function $f(z_1, z_2)$ *which is regular in* $\mathfrak{M}$ *can be represented in* $\mathfrak{M}$ *in the form*

$$f(t_1, t_2) = \frac{1}{2} \frac{1}{(2\pi i)^2} \sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k,s}^1} \frac{f(\varphi_1, \varphi_2) B_{k,s}(t_1, t_2, \lambda_k, \lambda_s) \, d\lambda_k \, d\lambda_s}{\Phi_k(t_1, t_2, \lambda_k) \Phi_s(t_1, t_2, \lambda_s)}, \tag{2.16}$$

where

$$B_{k,s} \equiv \frac{\Phi_s(t_1, t_2, \lambda_s) \Phi_k(t_1, \varphi_2, \lambda_k) - \Phi_k(t_1, t_2, \lambda_k) \Phi_s(t_1, \varphi_2, \lambda_s)}{(\varphi_1 - t_1)(\varphi_2 - t_2)} \frac{\partial (\varphi_1, \varphi_2)}{\partial (\lambda_k, \lambda_s)}. \tag{2.17}$$

[See (2.1) and (2.7)]. (2.16) expresses the value of the function f in $\mathfrak{M}$ in terms of the values on the distinguished boundary surface. It can be considered as a generalization of CAUCHY formula. When compared with the formula

$$\frac{1}{2\pi i} \int \frac{f(t)}{t - z(\lambda)} \frac{dz(\lambda)}{d\lambda} d\lambda,$$

(2.16) has the disadvantage that its kernel

$$\frac{B_{ks}(t_1, t_2, \lambda_k, \lambda_s)}{\Phi_k(t_1, t_2, \lambda_k)\,\Phi_s(t_1, t_2, \lambda_s)} \tag{2.18}$$

depends upon the domain, and therefore in many instances we cannot operate with it so easily as in the case of one variable.

In the second part of the present section we derive formulas of another type which express the values of a function in $\mathfrak{R}$ in terms of its values on the distinguished boundary surface.

§ 3. A representation of functions $f(z_1, z_2)$ in terms of their values on the distinguished boundary surface.

LEMMA 2.1. *With every point (t_1, t_2) of the domain $\mathfrak{R}$ we can associate bounded functions $M_{ks}(t_1, t_2, \lambda_k, \lambda_s)$ such that for $B_{ks}(t_1, t_2, \lambda_k, \lambda_s)$* [see (2.17)], *the inequality*

$$|B_{ks}(t_1, t_2, \lambda_k, \lambda_s)| \leq M_{ks}(t_1, t_2, \lambda_k, \lambda_s) < \infty, \qquad (t_1, t_2) \in \overline{\mathfrak{R}}, \tag{2.19}$$

holds.

PROOF. Let us consider at first the expression $\dfrac{\Phi_s(t_1, \varphi_2(\lambda_k, \lambda_s), \lambda_s)}{\varphi_1(\lambda_k, \lambda_s) - t_1}$ for a *fixed* λ_k, λ_s. $\Phi_s(t_1, t_2, \lambda_s)$ is defined for all values of $(t_1, t_2) \in \overline{\mathfrak{R}}$, $\Phi_s\big(t_1, \varphi_2(\lambda_k, \lambda_s), \lambda_s\big)$ is defined at the points t_1 such that $(t_1, t_2 = \varphi_2) \in \mathfrak{P}_1^2(\varphi_2)$, the intersection of $\overline{\mathfrak{R}}$ with the plane[4]) $t_2 = \varphi_2$, and vanishes for $t_1 = \varphi_1$.

$\Phi_2\big(t_1, \varphi_2(\lambda_k, \lambda_s), \lambda_s\big)$ is a regular function of t_1, therefore it must have a factor $t_1 - \varphi_1(\lambda_k, \lambda_s)$.

We consider

$$\frac{\dfrac{\Phi_k(t_1, \varphi_2, \lambda_k)}{\varphi_1(\lambda_k, \lambda_s) - t_1}\,\Phi_s(t_1, t_2, \lambda_s) - \dfrac{\Phi_s(t_1, \varphi_2, \lambda_s)}{\varphi_1(\lambda_k, \lambda_s) - t_1}\,\Phi_k(t_1, t_2, \lambda_k)}{\varphi_2(\lambda_k, \lambda_s) - t_2} \tag{2.20}$$

as a function of t_2, $(t_1 = \varphi_1, t_2) \in \overline{\mathfrak{P}_2^2}(\varphi_1)$. (See [4]). For $t_2 = \varphi_2(\lambda_k, \lambda_s)$ the numerator vanishes and therefore it must have a factor $t_2 - \varphi_2(\lambda_k, \lambda_s)$. Therefore, and in consequence of the hypothesis **2, c***, see p. 176, for $(t_1, t_2) \in \mathfrak{R}$ there exists an upper bound $M_{ks}(t_1, t_2, \lambda_k, \lambda_s)$ for $|B_{ks}(t_1, t_2, \lambda_k, \lambda_s)|$.

LEMMA 2.2. *A function $f(z_1, z_2)$ of two complex variables which is regular in $\mathfrak{R}$ and continuous in $\overline{\mathfrak{R}}$ satisfies the inequality*

$$|f(t_1, t_2)| \leq \frac{(n(n-1))^{\frac{1}{2}}}{2(2\pi)^2}\,[\mu(t_1, t_2)]^{\frac{1}{2}} \left[\sum_{k=1}^{n} \sum_{\substack{s=1 \\ k \neq s}}^{n} \iint_{\mathfrak{B}_k^1} |f(\varphi_1, \varphi_2)|^2 \left|\frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)}\right| d\lambda_k\, d\lambda_s\right]^{\frac{1}{2}}, \tag{2.21}$$

$$(t_1, t_2) \in \mathfrak{R},$$

where $\mu(t_1, t_2)$ is an upper bound for

$$\left\{\iint_{\mathfrak{B}_k^1} \frac{M_{ks}^2(t_1, t_2, \lambda_k, \lambda_s)}{|\Phi_k(t_1, t_2, \lambda_k)\,\Phi_s(t_1, t_2, \lambda_s)|^2} \left|\frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)}\right| d\lambda_k\, d\lambda_s\right\}, \tag{2.22}$$

$$k = 1, 2, \ldots, n, \qquad s = 1, 2, \ldots, n, \qquad k \neq s.$$

[4]) In a similar way we introduce the domain $\mathfrak{P}_2^2(\varphi_1)$ the intersection of $\overline{\mathfrak{R}}$ with the plane $t_1 = \varphi_1(\lambda_k, \lambda_s)$. See also (2.7).

PROOF. As it has been shown in [5], $f(t_1, t_2)$ can be represented in $\mathfrak{R}$ in the form (2.16). By SCHWARZ inequality and Lemma 2.1 follows from (2.16)

$$
(2.23)\quad \left\{
\begin{aligned}
|f(t_1,t_2)| &\leq \frac{1}{8\pi^2}\sum_{\substack{k=1\\k\neq s}}^{n}\sum_{s=1}^{n}\left[\left\{\iint_{\mathfrak{B}_k^1}|f(\varphi_1,\varphi_2)|^2\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\right\}^{\frac12}\times\right.\\
&\qquad\left.\times\left\{\iint_{\mathfrak{B}_k^1}\frac{M_{ks}^2(t_1,t_2,\lambda_k,\lambda_s)}{|\Phi_k(t_1,t_2,\lambda_k)\,\Phi_s(t_1,t_2,\lambda_s)|^2}\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\right\}^{\frac12}\right]\\
&\leq \frac{1}{8\pi^2}\left(\mu(t_1,t_2)\right)^{\frac12}\sum_{\substack{k=1\\k\neq s}}^{n}\sum_{s=1}^{n}\left[\iint_{\mathfrak{B}_k^1}|f(\varphi_1,\varphi_2)|^2\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\right]^{\frac12}\\
&\leq \frac{1}{8\pi^2}\left(\mu(t_1,t_2)\right)^{\frac12}\left(\frac{n(n-1)}{2}\right)^{\frac12}\left[\sum_{\substack{k=1\\k\neq s}}^{n}\sum_{s=1}^{n}\iint_{\mathfrak{B}_k^1}|f(\varphi_1,\varphi_2)|^2\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\right]^{\frac12}.
\end{aligned}
\right.
$$

REMARK. Deriving the last expression we use the fact that

$$
\left[\sum_{\varkappa=1}^{M}|a_\varkappa|^{\frac12}\right]^2=\sum_{\varkappa=1}^{M}|a_\varkappa|+\sum_{\substack{\varkappa=1\\\varkappa<\varrho}}^{M}\sum_{\varrho=1}^{M}2|a_\varkappa|^{\frac12}|a_\varrho|^{\frac12}\leq\sum_{\nu=1}^{M}|a_\nu|+\sum_{\nu=1}^{M-1}\left[\sum_{\varkappa=\nu}^{M}\left(|a_\varkappa|+|a_{\varkappa+\nu}|\right)\right]
$$

$$
=M\left(\sum_{\varkappa=1}^{M}|a_\varkappa|\right),\qquad M=\frac{n(n-1)}{2},
$$

THEOREM 2.1. *Let $\Phi_\nu(z_1,z_2)$, $\nu=1,2,\ldots$ be a set of functions which are regular in $\mathfrak{R}$, continuous in $\overline{\mathfrak{R}}$ and orthonormal in the sense that the integral over the distinguished boundary surface:*

$$
(2.24)\quad \left\{
\begin{aligned}
[\Phi_\nu,\overline{\Phi}_\mu]_{\mathfrak{D}^1}&\equiv\sum_{\substack{k=1\\k\neq s}}^{n}\sum_{s=1}^{n}\iint_{\mathfrak{B}_k^1}\Phi_\nu(\varphi_1,\varphi_2)\,\overline{\Phi_\mu(\varphi_1,\varphi_2)}\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\\
&=\delta_{\nu\mu}\begin{cases}=1 & \text{if}\quad \nu=\mu,\\ =0 & \text{if}\quad \nu\neq\mu.\end{cases}
\end{aligned}
\right.
$$

Then the kernel function

$$
(2.25)\qquad \hat{K}(z_1,z_2;\bar{\zeta}_1,\bar{\zeta}_2)=\sum_{\nu=1}^{\infty}\Phi_\nu(z_1,z_2)\,\overline{\Phi_\nu(\zeta_1,\zeta_2)}
$$

of this system converges absolutely at every interior point of the domain and uniformly in every closed subdomain of $\mathfrak{R}$.

PROOF. We set $\hat{K}_n(z_1,z_2;\bar{\zeta}_1,\bar{\zeta}_2)=\sum_{\nu=1}^{n}\Phi_\nu(z_1,z_2)\,\overline{\Phi_\nu(\zeta_1,\zeta_2)}$. We shall show that $\hat{K}_n(z_1,z_2;\bar{\zeta}_1,\bar{\zeta}_2)$ is uniformly bounded in every closed subdomain of $\mathfrak{R}$, independently of n. If $f(z_1,z_2)$ is analytic in $\overline{\mathfrak{R}}$ and if $(t_1,t_2)\in\mathfrak{R}$, then according to the lemma 2.2

$$
(2.26)\qquad \sum_{\substack{k=1\\k\neq s}}^{n}\sum_{s=1}^{n}\iint_{\mathfrak{B}_k^1}|f(\varphi_1,\varphi_2)|^2\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s\geq\delta(t_1,t_2)\,|f(t_1,t_2)|^2
$$

[see (2.7)] where

$$\delta(t_1, t_2) = \frac{2(8\pi^2)^2}{n(n-1)} \frac{1}{\mu(t_1, t_2)} > 0.$$

[Concerning μ see (2.21), (2.22).] From (2.26) we conclude that if $\{\Phi_\nu(z_1, z_2)\}$ is a system of functions which are regular in $\mathfrak{R}$ and orthonormalized in the sense of (2.24), then

$$\sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2 \leq [\delta(t_1, t_2)]^{-1}. \tag{2.27}$$

For choosing $f(z_1, z_2) = \sum_{\nu=1}^{N} \Phi_\nu(z_1, z_2)\, \overline{\Phi_\nu(t_1, t_2)}$ we obtain according to (2.23)

$$\left\{ \begin{aligned} &\left(\sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2 \right)^2 \\ &\leq \frac{n(n-1)}{2(8\pi^2)^2} \mu(t_1, t_2) \sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k,s}^1} \left| \sum_{\nu=1}^{N} \Phi_\nu(\varphi_1, \varphi_2)\, \overline{\Phi_\nu(t_1, t_2)} \right|^2 \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k\, d\lambda_s. \end{aligned} \right. \tag{2.28}$$

Hence using the orthogonality relations, see (2.24), we have

$$\left\{ \begin{aligned} &\frac{2(8\pi^2)^2}{n(n-1)} \frac{1}{\mu(t_1, t_2)} \left(\sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2 \right)^2 \\ &\leq \sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k,s}^1} \left| \sum_{\nu=1}^{N} \Phi_\nu(\varphi_1, \varphi_2)\, \overline{\Phi_\nu(t_1, t_2)} \right|^2 \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k\, d\lambda_s \\ &= \sum_{\nu=1}^{N} \sum_{\mu=1}^{N} \Phi_\nu(t_1, t_2)\, \overline{\Phi_\mu(t_1, t_2)} \sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k,s}^1} \overline{\Phi_\nu(\varphi_1, \varphi_2)}\, \Phi_\mu(\varphi_1, \varphi_2) \times \\ &\qquad \times \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k\, d\lambda_s = \sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2. \end{aligned} \right. \tag{2.29}$$

Since $\sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2 \neq 0$,

$$\frac{2(8\pi^2)^2}{n(n-1)} \frac{1}{\mu(t_1, t_2)} \sum_{\nu=1}^{N} |\Phi_\nu(t_1, t_2)|^2 \leq 1. \tag{2.30}$$

REMARK. It is clear that in addition to the orthogonal functions introduced in the present paper we can introduce functions $\psi_\nu(z_1, z_2)$, $\nu = 1, 2, \ldots$, which are orthonormal in the sense that

$$\left\{ \begin{aligned} \{\psi_\nu, \overline{\psi}_\mu\} \equiv &\sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k,s}^1} \psi_\nu(\varphi_1, \varphi_2)\, \overline{\psi_\mu(\varphi_1, \varphi_2)}\, W(\varphi_1, \varphi_2) \times \\ &\times \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k\, d\lambda_s = \delta_{\nu\mu}, \end{aligned} \right. \tag{2.31}$$

where $W(\varphi_1, \varphi_2)$ is a *positive* continuous weighting function defined on the distinguished boundary surface.

§ 4. A special set of orthonormal functions. We proceed now to a construction of a set of orthonormal functions each of which is regular in $\mathfrak{W}$. This set is complete for the class of functions which are regular in $\mathfrak{W}$.

We make the (additional) hypothesis **4**: There exist functions

$$\Phi_k(t_1, t_2, \lambda_k + i\,\lambda_k^*), \qquad k = 1, 2, \ldots, n,$$

regular for $(t_1, t_2) \in \overline{\mathfrak{W}}$ and continuous in $\lambda_k + i\,\lambda_k^*$, such that 1°.

$$\Phi_k(t_1, t_2, \lambda_k + i\,\lambda_k^*) = 0, \qquad k = 1, 2, \ldots, n, \tag{2.32}$$

for every fixed λ_k^*, $0 < \lambda_k^* \leq p$ ($p > 0$, sufficiently small) are analytic hypersurfaces which lie completely outside of $\mathfrak{W}$, 2°. for $\lambda_k^* = 0$ are identical with the functions considered on p. 175, see (2.1).

We determine on every surface $\mathfrak{B}_{ks}^2$ a set of points $(\lambda_k^\nu, \lambda_s^\nu)$, $\nu = 1, 2, \ldots$, which with increasing ν becomes denser and denser. Further for every N we determine in $\mathfrak{B}_{ks}^2$ surface segments $\mathfrak{A}_N^2(\lambda_k^\nu, \lambda_s^\nu)$, $\nu \leq N$, $\sum_{\nu=1}^{N} \mathfrak{A}_N^2(\lambda_k^\nu, \lambda_s^\nu) = \mathfrak{B}_{ks}^2$, such that the projection of every $\mathfrak{A}_N^2(\lambda_k^\nu, \lambda_s^\nu)$ on the λ_k, λ_s-plane includes in its interior the point $(\lambda_k^\nu, \lambda_s^\nu)$, but *do not* include any other point $(\lambda_k^\mu, \lambda_s^\mu)$, $\mu = 1, 2, \ldots, \nu-1, \nu+1, \ldots, N$. We denote by $\Delta_N(\lambda_\nu^k, \lambda_s^\nu)$ the area of the projection of $\mathfrak{A}_N^2(\lambda_k^\nu, \lambda_s^\nu)$ on the λ_k, λ_s-plane. The $(\lambda_k^\nu, \lambda_s^\nu)$ are chosen in such a way that $\lim_{N\to\infty} \Delta_N(\lambda_k^\nu, \lambda_s^\nu) = 0$, uniformly in ν.

Let $f(z_1, z_2)$ be a function of two complex variables which is regular in $\mathfrak{W}$. For $(t_1, t_2) \in \mathfrak{W}_1 < \mathfrak{W}$, every of the integrals on the right hand side of (2.16) can be uniformly approximated by

$$\sum_{\nu=1}^{N} \frac{f(\varphi_1^\nu, \varphi_2^\nu)\, B_{ks}(t_1, t_2, \lambda_k^\nu, \lambda_s^\nu)}{\Phi_k(t_1, t_2, \lambda_k^\nu)\, \Phi_s(t_1, t_2, \lambda_s^\nu)}\, \Delta_N(\lambda_k^\nu, \lambda_s^\nu) \tag{2.33}$$

where $(\lambda_k^\nu, \lambda_s^\nu)$, $\nu = 1, 2, \ldots$, is a sufficiently dense set of points on $\mathfrak{B}_{ks}^2$ and $\varphi_\varkappa^\nu = \varphi_\varkappa(\lambda_k^\nu, \lambda_s^\nu)$, $\varkappa = 1, 2$.

We approximate now the sum (2.33) by

$$\left\{ \sum_{\nu=1}^{N} \frac{f(\varphi_1^\nu, \varphi_2^\nu)\, B_{ks}(t_1, t_2, \lambda_k^\nu, \lambda_s^\nu)}{\Phi_k(t_1, t_2, \lambda_k^\nu + i\,2^{-\nu}p)\, \Phi_s(t_1, t_2, \lambda_s^\nu + i\,2^{-\nu}p)}\, \Delta_N(\lambda_k^\nu, \lambda_s^\nu), \quad p > 0, \right. \tag{2.34}$$

small enough.

By the hypothesis **4** and according to the considerations of [5] every term

$$\frac{B_{ks}(t_1, t_2, \lambda_k^\nu, \lambda_s^\nu)}{\Phi_k(t_1, t_2, \lambda_k^\nu + i\,2^{-\nu}p)\, \Phi_s(t_1, t_2, \lambda_s + i\,2^{-\nu}p)}, \qquad p > 0, \tag{2.35}$$

is a regular function for $(t_1, t_2) \in \mathfrak{R}$. Since the $B_{ks}(t_1, t_2, \lambda_k^\nu, \lambda_s^\nu)$, $\nu = 1, 2, \ldots, N$, are regular for $(t_1, t_2) \in \mathfrak{R}$ it suffices to show that

$$(2.36) \qquad \left\{ \sum_{\nu=1}^{N} |[\Phi_k(t_1, t_2, \lambda_k^\nu + i\,2^{-\nu} p)\, \Phi_s(t_1, t_2, \lambda_s^\nu + i\,2^{-\nu} p)]^{-1} - [\Phi_k(t_1, t_2, \lambda_k^\nu)\, \Phi_s(t_1, t_2, \lambda_s^\nu)]^{-1}| \, \Delta_N(\lambda_k^\nu, \lambda_s^\nu) \right.$$

goes to zero for $N \to \infty$, and $(t_1, t_2) \in \mathfrak{R}_1 < \mathfrak{R}$.

As we mentioned before with increasing N the $\Delta_N(\lambda_s^\nu, \lambda_k^\nu)$ go to zero (uniformly in ν). On the other hand for every fixed ε, there exists only a finite number of points $(\lambda_k^\nu, \lambda_s^\nu)$, say $\nu = 1, 2, \ldots, T$, for which

$$(2.37) \qquad \left\{ |[\Phi_k(t_1, t_2, \lambda_k^\nu + i\,2^{-\nu} p)\, \Phi_s(t_1, t_2, \lambda_s^\nu + i\,2^{-\nu} p)]^{-1} - [\Phi_k(t_1, t_2, \lambda_k^\nu)\, \Phi_s(t_1, t_2, \lambda_s^\nu)]^{-1}| > \varepsilon. \right.$$

If we denote by m the maximum of the left hand side of (2.37) for $(t_1, t_2) \in \mathfrak{R}_1$ then (2.36) will be smaller than

$$(2.38) \qquad m\left[\sum_{\nu=1}^{T} \Delta_N(\lambda_k^\nu, \lambda_s^\nu)\right] + \varepsilon A(\mathfrak{B}_{ks}^2)$$

where $A(\mathfrak{B}_{ks}^2)$ denotes the area of the projection of $\mathfrak{B}_{ks}^2$ on the λ_k, λ_s-plane.

(Note that T is *independent* of N.) With increasing N, $\sum_{\nu=1}^{T} \Delta_N(\lambda_k^\nu, \lambda_s^\nu) \to 0$ and therefore for every ε we can determine an N so that (2.38) becomes smaller than $2\varepsilon A(\mathfrak{B}_{ks}^2)$. Since ε can be chosen arbitrarily small, (2.36) goes to zero for $\varepsilon \to 0$.

Thus we obtain the

THEOREM 2.2. *By orthogonalizing in the sense* (2.24) *the functions*

$$(2.39) \qquad \left\{ \frac{B_{ks}(t_1, t_2, \lambda_k^\nu, \lambda_s^\nu)}{\Phi_k(t_1, t_2, \lambda_k^\nu + i\,2^{-\nu} p)\, \Phi_s(t_1, t_2, \lambda_s^\nu + i\,2^{-\nu} p)} \quad p > 0, \text{ small enough,} \right.$$
$$\nu = 1, 2, 3, \ldots, \; k = 1, 2, \ldots, n, \quad s = 1, 2, \ldots, n, \quad s \neq k,$$

we obtain a system of orthonormal functions each of which is regular in $\mathfrak{R}$. *This system is complete for the class of functions* $f(t_1, t_2)$ *which are regular in* $\mathfrak{R}$.

For every domain $\mathfrak{R}$ which admits a group of transformations

$$(2.40) \qquad \left\{ \begin{array}{l} z_1^* = q_1(z_1, z_2, \varepsilon), \qquad z_2^* = q_2(z_1, z_2, \varepsilon), \qquad 0 < \varepsilon < \varepsilon_0, \\ q_\varkappa(z_1, z_2, 0) = z_\varkappa, \qquad \varkappa = 1, 2, \end{array} \right.$$

which transforms $\mathfrak{R}$ into a domain $\mathfrak{R}_\varepsilon > \mathfrak{R} \equiv \mathfrak{R}_0$ (e.g. for every convex domain) the theorems which were proved for the class of functions which are regular in $\mathfrak{R}$, hold also for functions which have on the distinguished boundary surface $\mathfrak{D}^2$ [see (2.6a)] boundary values almost everywhere, and for which $[f, \bar{f}] < \infty$. We shall denote this class by $\tilde{L}^2(\mathfrak{R})$. Indeed, if we replace $f(z_1, z_2)$ by $f_\varepsilon(z_1, z_2) = f[q_1(z_1, z_2, \varepsilon), q_2(z_1, z_2, \varepsilon)]$ for $\varepsilon > 0$, f_ε is regular in $\mathfrak{R}$. Since $\lim_{\varepsilon \to 0} [f_\varepsilon, \bar{f}_\varepsilon] = [f, \bar{f}]$ our assertion follows. Here $[\,,\,] \equiv [\,,\,]_{\mathfrak{D}^2}$.

§ 5. An application to the theory of meromorphic functions.

THEOREM 2.3. *Let $g_\nu(t_1, t_2)$, $\nu = 1, 2, \ldots, m$, be a (finite) set of functions which are regular in $\mathfrak{R}$ and orthonormalized in the sense* (2.24). *Suppose that the function $f(t_1, t_2)$, $f \in \tilde{L}^2(\mathfrak{R})$ is orthogonal to the $g_\nu(t_1, t_2)$, i.e.*[4a]) $[f, \bar{g}_\nu] = 0$ *for* $\nu = 1, 2, \ldots, m$. *Then*

$$|f(t_1, t_2)|^2 \leq [f, \bar{f}] \left(\hat{K}(t_1, t_2; \bar{t}_1, \bar{t}_2) - \sum_{\nu=1}^{m} |g_\nu(t_1, t_2)|^2 \right). \tag{2.41}$$

PROOF. Since the g_ν, $\nu = 1, 2, \ldots, m$, are orthonormal we can determine a system $\{\chi_\mu\}$, $\mu = 1, 2, \ldots$ of orthonormal functions, such that the χ_μ are orthonormal to g_ν, i.e. that $[g_\nu, \bar{\chi}_\mu] = 0$ for every ν and μ, and such that $(\{g_\nu\} + \{\chi_\mu\})$ forms a set complete for the functions of the class $\tilde{L}^2(\mathfrak{R})$. Then it holds

$$f(t_1, t_2) = \sum_{\nu=1}^{m} [f, \bar{g}_\nu]\, g_\nu(t_1, t_2) + \sum_{\nu=1}^{\infty} [f, \bar{\chi}_\nu]\, \chi_\nu(t_1, t_2) = \sum_{\nu=1}^{\infty} [f, \bar{\chi}_\nu]\, \chi_\nu(t_1, t_2). \tag{2.42}$$

Since

$$\sum_{\nu=1}^{m} |g_\nu(t_1, t_2)|^2 + \sum_{\nu=1}^{\infty} |\chi_\nu(t_1, t_2)|^2 = \hat{K}(t_1, t_2; \bar{t}_1, \bar{t}_2)$$

the inequality (2.41) follows.

Let $M(t_1, t_2)$ be a meromorphic function of t_1, t_2 in $\mathfrak{R}$, which possesses the boundary values on $\mathfrak{D}^2$ almost everywhere and such that $[M, \bar{M}] < \infty$. Let $\{\varphi_\nu(t_1, t_2)\}$ be a complete set of orthonormal functions, which are regular in $\mathfrak{R}$.

Let further

$$m(t_1, t_2) = M(t_1, t_2) - \sum_{\nu=1}^{\infty} \varphi_\nu(t_1, t_2)\, [M, \bar{\varphi}_\nu]. \tag{2.43}$$

Then the (meromorphic) function $m(t_1, t_2)$ has the property that

$$[m, \bar{f}] = 0 \tag{2.44}$$

for every function $f \in \tilde{L}^2(\mathfrak{R})$.

Indeed for every function $\varphi_\mu(t_1, t_2)$ of our system, we have

$$\left[\left(M(t_1, t_2) - \sum_{\nu=1}^{\infty} \varphi_\nu(t_1, t_2)\, [M, \bar{\varphi}_\nu] \right), \overline{\varphi_\mu(t_1, t_2)} \right] = [M, \bar{\varphi}_\mu] - [M, \bar{\varphi}_\mu] = 0. \tag{2.45}$$

Since for $f = \sum_{\nu=1}^{\infty} a_\nu \varphi_\nu \in \tilde{L}^2(\mathfrak{R})$ we can integrate term by term, the relation (2.44) follows.

We shall call the functions m, satisfying for every $f \in \tilde{L}^2(\mathfrak{R})$ the relation (2.44) as *normalized with respect to the class* $\tilde{L}^2(\mathfrak{R})$.

[4a]) In the following *in this section* we omit the subscript $\mathfrak{D}^2$ in $[\,,\,]_{\mathfrak{D}^2}$.

We have proved:

THEOREM 2.4. *Let M be a function which is meromorphic in $\mathfrak{R}$, and m the corresponding function normalised with respect to $\tilde{L}^2(\mathfrak{R})$. If $f=M-m$, see (2.43), then*

$$[M,\overline{M}]=[m,\overline{m}]+[f,\overline{f}]. \tag{2.46}$$

§ 6. The method of the minimum integral in the case of the metric (2.24). The method based on the use of orthonormal systems seems to gain importance in recent investigations in the theory of functions of several complex variables[5]). The object of this section is to show how the method of the minimum integral (see [7], chapter III) has to be modified for functions which are orthonormal with respect to metric (2.24) (*orthonormalization over the distinguished boundary surface*). In this connection it is useful to employ a slightly different class of functions than those considered in § 3 [see (2.24)]. In order to do this we have to limit ourselves to a special class of domains which we shall denote by $\mathfrak{R}_I$ and which, in addition to the hypotheses **1**, **2**, **3** formulated in § 2 (see also [4, 5]), satisfy the following hypotheses.

5.1. There exists a sequence of domains $\mathfrak{R}_\varkappa$, $\varkappa=1,2,\ldots$, $\overline{\mathfrak{R}}_\varkappa<\mathfrak{R}_{\varkappa+1}$, each of which has the distinguished boundary surface $\mathfrak{D}^2_\varkappa$ and satisfies the hypotheses **1**, **2**, **3**. Each $\mathfrak{D}^2_\varkappa=\sum_{k=1}^{n}\sum_{\substack{s=1\\k\neq s}}^{n}\mathfrak{G}^2_{k,s,\varkappa}$ has the same structure as $\mathfrak{D}^2$ (see § 2). In particular, each $\mathfrak{G}^2_{k,s,\varkappa}=\mathfrak{S}\mathfrak{R}^2_{k,s,\varkappa\mu}$ where $\mathfrak{R}^2_{k,s,\varkappa\mu}$ can be represented in the form $z_m=\varphi_{m,\varkappa}(\lambda_k,\lambda_s)$, $m=1,2$. The $\varphi_{m,\varkappa}(\lambda_k,\lambda_s)$ converge uniformly to $\varphi_m(\lambda_k,\lambda_s)$. We assume further that $\mathfrak{R}_\varkappa$ and $\mathfrak{G}^2_{k,s,\varkappa}$ converge (in the sense of FRÉCHET) to $\mathfrak{R}_I$ and $\mathfrak{G}^2_{k,s,I}$, respectively.

5.2. We associate with every domain $\mathfrak{R}_\varkappa$ a function $\mu_\varkappa(t_1,t_2)$ introduced in lemma 2.2 [see (2.21)] and we assume that in every subdomain $\overline{\mathfrak{E}}$ which lies completely inside of $\mathfrak{R}$, the $\mu_\varkappa(t_1,t_2)$, $\varkappa=1,2,\ldots$ are uniformly bounded, i.e., to every $\overline{\mathfrak{E}}$ there exists a constant $C_1=C_1(\mathfrak{E})$ such that

$$\frac{n(n-1)}{128\pi^4}\mu_\varkappa(t_1,t_2)\leq C_1,\qquad (t_1,t_2)\in\overline{\mathfrak{E}},\qquad \varkappa=1,2,\ldots \tag{2.47}$$

[see (2.23)].

DEFINITION. A function $f(z_1,z_2)$, $(z_1,z_2)\in\mathfrak{R}_I$, which is regular in $\mathfrak{R}_I$ will be said to belong to the class $\hat{L}^2(\mathfrak{R}_I)$ if

$$\{f,\overline{f}\}_{\mathfrak{D}_I^2}\equiv\lim_{\varkappa\to\infty}\sum_{k=1}^{n}\sum_{\substack{s=1\\k\neq s}}^{n}\iint_{\mathfrak{B}^2_{k,s,\varkappa}}|f(\varphi_{1,\varkappa},\varphi_{2,\varkappa})|^2\left|\frac{\partial(\varphi_{1,\varkappa},\varphi_{2,\varkappa})}{\partial(\lambda_k,\lambda_s)}\right|d\lambda_k\,d\lambda_s<\infty. \tag{2.48}$$

[5]) See, e.g., BERGMAN [1, 6, 7, 8, 9], BREMERMANN [12], DAVIS [13], MASCHLER [14a], NEHARI [16]. In a number of papers (see [15]) and publications mentioned there Mr. H. MESCHKOWSKI applied the methods based on the theory of orthogonal functions for the study of analytic functions of one complex variable. It is of considerable interest that many of his results (with the necessary modifications) hold in the case of functions of two (and more) complex variables, in particular when these functions are considered in domains with a distinguished boundary surface.

THEOREM 2.5. *Let $\mathfrak{R}_I$ be a domain which satisfies hypotheses* **1–3, 5**. *Further, let $\mathfrak{R}$ be a domain with the distinguished boundary surface $\mathfrak{D}^2$ whose closure lies completely inside $\mathfrak{R}_I$.*

Then for every $f \in \hat{L}^2(\mathfrak{R}_I)$ the inequality

$$[f, \bar{f}]_{\mathfrak{D}^2} \leq C_2 A \{f, \bar{f}\}_{\mathfrak{D}_I^2} \tag{2.49}$$

holds [see (2.24) and (2.48)]. *Here A is the "area" of the projection of $\mathfrak{D}^2$, and C_2 is a constant depending only on $\mathfrak{R}$ and $\mathfrak{R}_I$.*

PROOF. Let $\mathfrak{R}_\varkappa$, $\varkappa = 1, 2, \ldots$ be a sequence of domains converging in the sense indicated in **5.1** to $\mathfrak{R}_I$. Let us consider $f(z_1, z_2)$ in the domain $\mathfrak{R}_\varkappa$, $\varkappa < \infty$. The function $f(z_1, z_2)$ is regular in $\mathfrak{R}_\varkappa$ and therefore according to (2.23) at every point $(t_1, t_2) \in \mathfrak{D}^2$ the inequality

$$|f(t_1, t_2)|^2 \leq C \mu_\varkappa(t_1, t_2) \left[\sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k s, \varkappa}^1} |f(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})|^2 \left| \frac{\partial(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k \, d\lambda_s \right], \tag{2.50}$$

$$C \equiv n(n-1)/128\pi^4.$$

holds for $\varkappa$ large enough. Since for $\varkappa$ large enough $\overline{\mathfrak{R}} \subset \mathfrak{R}_\varkappa \subset \mathfrak{R}_I$

$$|f(t_1, t_2)|^2 \leq C \mu_\varkappa [f, \bar{f}]_{\mathfrak{D}_\varkappa^2} \leq C(\mu + \varepsilon) \{f, \bar{f}\}_{\mathfrak{D}_I^2}. \tag{2.51}$$

Since ε can be made arbitrarily small, by integrating (2.51) over $\mathfrak{D}^2$, (2.49) follows.

When considering functions which are orthonormal when integrating over the domain, i.e., functions $\Phi_\nu(z_1, z_2)$ for which

$$\iiiint_{\mathfrak{R}} \Phi_\nu \bar{\Phi}_\mu \, dx_1 \, dy_1 \, dx_2 \, dy_2 = \delta_{\nu\mu}, \qquad z_k = x_k + i y_k, \qquad k = 1, 2, \tag{2.52}$$

we connected the theory of these functions with the method of the minimum integral (see, in particular, [7] p. 40ff.). Most of these considerations are rather of a formal nature and can immediately be repeated in the case of functions which are orthonormal in the sense of the metric (2.48). However, a modification of the procedure is needed. In the case of the metric (2.52), if we consider the minimum $\lambda_{\mathfrak{R}}$ of $\iiiint_{\mathfrak{R}} |f|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2$ for functions f, satisfying at a point $(t_1, t_2) \in \mathfrak{R} \in \mathfrak{R}_1$ certain conditions, the inequality

$$\begin{cases} \lambda_{\mathfrak{R}} \equiv \iiiint_{\mathfrak{R}} |f_{\mathfrak{R}}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \leq \iiiint_{\mathfrak{R}} |f_{\mathfrak{R}_1}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \\ \qquad\qquad \leq \iiiint_{\mathfrak{R}_1} |f_{\mathfrak{R}_1}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \equiv \lambda_{\mathfrak{R}_1} \end{cases} \tag{2.53}$$

holds (see [7], p. 44). Here $f_{\mathfrak{R}}$ (and $f_{\mathfrak{R}_1}$) is the function which minimizes the above integral, under conditions specified in [7] pp. 41ff. $\lambda_{\mathfrak{R}}$ depends monotonically upon the domain $\mathfrak{R}$.

In the case of orthonormal functions considered in the present section we have to replace the inequality $\lambda_{\mathfrak{R}} \leq \lambda_{\mathfrak{R}_1}$ for $\mathfrak{R} \subset \mathfrak{R}_1$ [see (2.53)] by the inequality (2.49). If in particular, we then choose for $\mathfrak{D}^2$, the distinguished

We have proved:

THEOREM 2.4. *Let M be a function which is meromorphic in $\mathfrak{R}$, and m the corresponding function normalised with respect to $\tilde{L}^2(\mathfrak{R})$. If $f = M - m$, see (2.43), then*

$$[M, \overline{M}] = [m, \overline{m}] + [f, \overline{f}]. \tag{2.46}$$

§ 6. The method of the minimum integral in the case of the metric (2.24). The method based on the use of orthonormal systems seems to gain importance in recent investigations in the theory of functions of several complex variables[5]). The object of this section is to show how the method of the minimum integral (see [7], chapter III) has to be modified for functions which are orthonormal with respect to metric (2.24) (*orthonormalization over the distinguished boundary surface*). In this connection it is useful to employ a slightly different class of functions than those considered in § 3 [see (2.24)]. In order to do this we have to limit ourselves to a special class of domains which we shall denote by $\mathfrak{R}_I$ and which, in addition to the hypotheses **1**, **2**, **3** formulated in § 2 (see also [*4*, *5*]), satisfy the following hypotheses.

5.1. There exists a sequence of domains $\mathfrak{R}_\varkappa$, $\varkappa = 1, 2, \ldots$, $\overline{\mathfrak{R}}_\varkappa < \mathfrak{R}_{\varkappa+1}$, each of which has the distinguished boundary surface $\mathfrak{D}_\varkappa^2$ and satisfies the hypotheses **1**, **2**, **3**. Each $\mathfrak{D}_\varkappa^2 = \sum_{k=1}^{n} \sum_{\substack{s=1 \\ k \neq s}}^{n} \mathfrak{G}_{ks,\varkappa}^2$ has the same structure as $\mathfrak{D}^2$ (see § 2). In particular, each $\mathfrak{G}_{ks,\varkappa}^2 = \sum \mathfrak{R}_{ks,\varkappa\mu}^2$ where $\mathfrak{R}_{ks,\varkappa\mu}^2$ can be represented in the form $z_m = \varphi_{m,\varkappa}(\lambda_k, \lambda_s)$, $m = 1, 2$. The $\varphi_{m,\varkappa}(\lambda_k, \lambda_s)$ converge uniformly to $\varphi_m(\lambda_k, \lambda_s)$. We assume further that $\mathfrak{R}_\varkappa$ and $\mathfrak{G}_{ks,\varkappa}^2$ converge (in the sense of FRÉCHET) to $\mathfrak{R}_I$ and $\mathfrak{G}_{ks,I}^2$, respectively.

5.2. We associate with every domain $\mathfrak{R}_\varkappa$ a function $\mu_\varkappa(t_1, t_2)$ introduced in lemma 2.2 [see (2.21)] and we assume that in every subdomain $\overline{\mathfrak{E}}$ which lies completely inside of $\mathfrak{R}$, the $\mu_\varkappa(t_1, t_2)$, $\varkappa = 1, 2, \ldots$ are uniformly bounded, i.e., to every $\overline{\mathfrak{E}}$ there exists a constant $C_1 = C_1(\mathfrak{E})$ such that

$$\frac{n(n-1)}{128\pi^4}\,\mu_\varkappa(t_1, t_2) \leq C_1, \qquad (t_1, t_2) \in \overline{\mathfrak{E}}, \qquad \varkappa = 1, 2, \ldots \tag{2.47}$$

[see (2.23)].

DEFINITION. A function $f(z_1, z_2)$, $(z_1, z_2) \in \mathfrak{R}_I$, which is regular in $\mathfrak{R}_I$ will be said to belong to the class $\hat{L}^2(\mathfrak{R}_I)$ if

$$\{f, \bar{f}\}_{\mathfrak{D}_I^2} \equiv \lim_{\varkappa \to \infty} \sum_{k=1}^{n} \sum_{\substack{s=1 \\ k \neq s}}^{n} \iint_{\mathfrak{B}_{ks,\varkappa}^2} |f(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})|^2 \left| \frac{\partial(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k\, d\lambda_s < \infty. \tag{2.48}$$

[5]) See, e.g., BERGMAN [*1*, *6*, *7*, *8*, *9*], BREMERMANN [*12*], DAVIS [*13*], MASCHLER [*14a*], NEHARI [*16*]. In a number of papers (see [*15*]) and publications mentioned there Mr. H. MESCHKOWSKI applied the methods based on the theory of orthogonal functions for the study of analytic functions of one complex variable. It is of considerable interest that many of his results (with the necessary modifications) hold in the case of functions of two (and more) complex variables, in particular when these functions are considered in domains with a distinguished boundary surface.

THEOREM 2.5. *Let $\mathfrak{M}_I$ be a domain which satisfies hypotheses* **1–3, 5**. *Further, let $\mathfrak{M}$ be a domain with the distinguished boundary surface $\mathfrak{D}^2$ whose closure lies completely inside $\mathfrak{M}_I$.*

Then for every $f \in \widehat{L}^2(\mathfrak{M}_I)$ the inequality

$$[f, \bar{f}]_{\mathfrak{D}^2} \leq C_2 A \{f, \bar{f}\}_{\mathfrak{D}_I^2} \tag{2.49}$$

holds [see (2.24) and (2.48)]. *Here A is the "area" of the projection of $\mathfrak{D}^2$, and C_2 is a constant depending only on $\mathfrak{M}$ and $\mathfrak{M}_I$.*

PROOF. Let $\mathfrak{M}_\varkappa$, $\varkappa = 1, 2, \ldots$ be a sequence of domains converging in the sense indicated in **5.1** to $\mathfrak{M}_I$. Let us consider $f(z_1, z_2)$ in the domain $\mathfrak{M}_\varkappa$, $\varkappa < \infty$. The function $f(z_1, z_2)$ is regular in $\mathfrak{M}_\varkappa$ and therefore according to (2.23) at every point $(t_1, t_2) \in \mathfrak{D}^2$ the inequality

$$|f(t_1, t_2)|^2 \leq C \mu_\varkappa(t_1, t_2) \left[\sum_{\substack{k=1 \\ k \neq s}}^{n} \sum_{s=1}^{n} \iint_{\mathfrak{B}_{k\,s,\varkappa}^1} |f(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})|^2 \left| \frac{\partial(\varphi_{1,\varkappa}, \varphi_{2,\varkappa})}{\partial(\lambda_k, \lambda_s)} \right| d\lambda_k \, d\lambda_s \right], \tag{2.50}$$

$$C \equiv n(n-1)/128\pi^4.$$

holds for $\varkappa$ large enough. Since for $\varkappa$ large enough $\overline{\mathfrak{M}} < \mathfrak{M}_\varkappa < \mathfrak{M}_I$

$$|f(t_1, t_2)|^2 \leq C \mu_\varkappa [f, \bar{f}]_{\mathfrak{D}_\varkappa^2} \leq C(\mu + \varepsilon) \{f, \bar{f}\}_{\mathfrak{D}_I^2}. \tag{2.51}$$

Since ε can be made arbitrarily small, by integrating (2.51) over $\mathfrak{D}^2$, (2.49) follows.

When considering functions which are orthonormal when integrating over the domain, i.e., functions $\Phi_\nu(z_1, z_2)$ for which

$$\iiiint_{\mathfrak{M}} \Phi_\nu \bar{\Phi}_\mu \, dx_1 \, dy_1 \, dx_2 \, dy_2 = \delta_{\nu\mu}, \qquad z_k = x_k + i y_k, \qquad k = 1, 2, \tag{2.52}$$

we connected the theory of these functions with the method of the minimum integral (see, in particular, [7] p. 40ff.). Most of these considerations are rather of a formal nature and can immediately be repeated in the case of functions which are orthonormal in the sense of the metric (2.48). However, a modification of the procedure is needed. In the case of the metric (2.52), if we consider the minimum $\lambda_{\mathfrak{M}}$ of $\iiiint_{\mathfrak{M}} |f|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2$ for functions f, satisfying at a point $(t_1, t_2) \in \mathfrak{M} \in \mathfrak{M}_1$ certain conditions, the inequality

$$\begin{cases} \lambda_{\mathfrak{M}} \equiv \iiiint_{\mathfrak{M}} |f_{\mathfrak{M}}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \leq \iiiint_{\mathfrak{M}} |f_{\mathfrak{M}_1}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \\ \qquad\qquad \leq \iiiint_{\mathfrak{M}_1} |f_{\mathfrak{M}_1}|^2 \, dx_1 \, dy_1 \, dx_2 \, dy_2 \equiv \lambda_{\mathfrak{M}_1} \end{cases} \tag{2.53}$$

holds (see [7], p. 44). Here $f_{\mathfrak{M}}$ (and $f_{\mathfrak{M}_1}$) is the function which minimizes the above integral, under conditions specified in [7] pp. 41ff. $\lambda_{\mathfrak{M}}$ depends monotonically upon the domain $\mathfrak{M}$.

In the case of orthonormal functions considered in the present section we have to replace the inequality $\lambda_{\mathfrak{M}} \leq \lambda_{\mathfrak{M}_1}$ for $\mathfrak{M} < \mathfrak{M}_1$ [see (2.53)] by the inequality (2.49). If in particular, we then choose for $\mathfrak{D}^2$, the distinguished

boundary surface of a bicylinder, whose closure lies in $\mathfrak{R}$, $\mathfrak{D}^2=[z_k=\varrho\exp(i\varphi_k)$, $k=1,2]$ and for which the kernel function in an explicit form is known, using (2.49) we can determine lower bounds for $\{f,\bar{f}\}_{\mathfrak{D}_1^2}$ which are positive.

In a similar manner we can determine upper bounds for the $\lambda_{\mathfrak{M}_1}$.

Thus most of the results which one can obtain by using the method of the minimum integral hold in a modified form also in the case of the metric (2.49).

III. The relations between the value distribution of a function $f(z_1, z_2)$ in the domain $\mathfrak{R}$ and on its boundary $\mathfrak{m}^3$.

§ 1. Bounds for $|f(z_1, z_2)|$ in an analytic surface which intersect the boundary hypersurface $\mathfrak{i}_k^3$ along a closed curve. In the previous section we discussed the connection between the behavior of a function on the distinguished boundary surface and in the interior of a domain $\mathfrak{R}$ possessing such a surface. The boundary $\mathfrak{m}^3$ of the (*four-dimensional*) domain $\mathfrak{R}$ consists of the (*two-dimensional*) distinguished boundary surface $\mathfrak{D}^2$ and of the (*three-dimensional*) complementary part $\mathfrak{b}^3=\mathfrak{m}^3-\mathfrak{D}^2$. There arises a problem to investigate the relations between the behavior of the function f inside the domain $\mathfrak{R}$ and in $\mathfrak{b}^3$. It is of interest that relations of this type have in many instances a completely different character than those obtained in the case of one variable, when investigating the relations between the behavior of a function $f(z)$ (of one variable) inside a domain and on its boundary curve. We proceed now to the description of certain properties of segments of the boundary $\mathfrak{b}^3$, which properties play a role for the purposes mentioned above.

As we indicated in § II, 2, $\mathfrak{m}^3=\sum_{k=1}^{n}\mathfrak{i}_k^3$ where

$$\mathfrak{i}_k^3=\sum_{\lambda_k\in\mathfrak{s}^1}\overline{\mathfrak{I}_k^2}(\lambda_k),\quad \mathfrak{s}^1=[0\leq\lambda_k\leq 2\pi], \tag{3.1}$$

whereby

$$\overline{\mathfrak{I}_k^2}(\lambda_k)=[z_\nu=h_\nu^k(Z_k,\lambda_k),\ \nu=1,2]. \tag{3.1a}$$

Let us assume that every $\overline{\mathfrak{I}_k^2}(\lambda_k)$ is a simply-connected (closed) domain and for every λ_k, $0\leq\lambda_k\leq 2\pi$, the point $A_k(\lambda_k)=[z_\nu=h_\nu^k(0,\lambda_k),\nu=1,2]$ corresponds to the origin $Z_k=0$. We assume further that the representation (3.1a) of the $\overline{\mathfrak{I}_k^2}(\lambda_k)$ is arranged in such a way that for every λ_k, $0\leq\lambda_k<2\pi$, Z_k varies in $|Z_k|\leq 1$. In this case we shall say that the representation of the segment $\mathfrak{i}_k^3$ is *normalized with respect to the axis* $\mathfrak{k}_k^1=\sum_{\lambda_k\in\mathfrak{s}^1}A_k(\lambda_k)$. Let us assume further that a closed curve

$$\mathfrak{p}_k^1=\sum_{\lambda_k\in\mathfrak{s}^1}P_k(\lambda_k),\qquad P_k(\lambda_k)=[z_\nu=h_\nu^k(\tau_k(\lambda_k),\lambda_k),\ \nu=1,2] \tag{3.2}$$

[where $\tau_k(\lambda_k)$, $0\leq\lambda_k\leq 2\pi$ is a continuous (single-valued) function of λ_k] is given[6]) on $\mathfrak{i}_k^3$. If $|\tau_k(\lambda_k)|\leq r$, $0\leq\lambda_k<2\pi$, $r<1$, the curve $\mathfrak{p}_k^1$ lies in the segment

[6]) Here we assume, that also $\mathfrak{I}_k^2(0)=\mathfrak{I}_k^2(2\pi)$ and therefore $h_\nu^k(Z_k,0)=h_\nu^k(Z_k,2\pi)$, $\nu=1,2$.

$\mathfrak{t}_k^3 < \mathfrak{i}_k^3$ whose representation can be written

$$\mathfrak{t}_k^3 = [z_\nu = h_\nu^k(Z_k, \lambda_k),\ |Z_k| < r,\ \nu = 1, 2]. \tag{3.3}$$

In this case we shall say that $\mathfrak{t}_k^3$ has *a representation normalized with respect to the axis* $\mathfrak{t}_k^1$ and in this representation it is of *radius* r.

THEOREM 3.1. *Suppose that the intersection of an analytic surface* $\mathfrak{A}_0^2 \equiv [z_\nu = g_\nu(\zeta), \nu = 1, 2]$ *with the boundary* $\mathfrak{m}^3$ *of* $\mathfrak{M}$ *is a closed curve* $\mathfrak{p}_k^1 = [z_\nu = g_\nu(\exp(i\varphi)),\ 0 \leq \varphi \leq 2\pi]$ *which lies in the segment* $\mathfrak{i}_k^3$.

Let us further make the following assumptions:

1. $\mathfrak{p}_k^1$ *lies in a segment* $\mathfrak{t}_k^3$ *which when its representation is normalized with respect to the axis* $\mathfrak{t}_k^1$ *is of radius* r, $r < 1$.

2. *An analytic function* $f(z_1, z_2)$ *is given which is regular in* $\overline{\mathfrak{M}}$, *and omits in* $\overline{\mathfrak{i}_k^3}$ *the values* 0 *and* 1.

3. *Along the axis* $\mathfrak{t}_k^1$ *of* $\mathfrak{t}_k^3$, f *assumes values which belong to a bounded (closed) domain* $\overline{\mathfrak{P}_0^2}$, *such that the points* (0), (1) $\notin \overline{\mathfrak{P}_0^2}$.

Then in $\mathfrak{A}^2 = \mathfrak{A}_0^2 \cap \mathfrak{M}$

$$|f(z_1, z_2)| \leq \exp[A(\mathfrak{P}_0^2)\,(1 - r)^{-1}] \tag{3.4}$$

where $A(\mathfrak{P}_0^2)$ *is a "universal" positive constant which is finite for every domain* $\mathfrak{P}_0^2$.

PROOF. As we indicated in § 2, $\mathfrak{i}_k^3$ is the sum of segments $\mathfrak{J}_k^2(\lambda_k)$ of analytic surfaces. In every segment $\overline{\mathfrak{J}_k^2(\lambda_k)}$ the function $F_k(Z_k, \lambda_k) \equiv f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$, $|Z_k| \leq 1$, for $0 \leq \lambda_k \leq 2\pi$, considered as a function of Z_k assumes at the point $Z_k = 0$ values belonging to $\overline{\mathfrak{P}_0^2}$; in the circle $|Z_k| \leq 1$, $F_k(Z_k, \lambda_k)$ omits the values 0 and 1. Consequently by the theorem of SCHOTTKY (see [*16*] p. 198) we have

$$f[h_1^k(Z_k, \lambda_k),\ h_2^k(Z_k, \lambda_k)] \leq \exp[A(\mathfrak{P}_0^2)\,(1 - r)^{-1}]. \tag{3.5}$$

According to our hypothesis, the boundary $\mathfrak{p}_k^1$ of the segment $\mathfrak{A}^2 = [z_1 = g_1(\zeta), z_2 = g_2(\zeta), |\zeta| \leq 1]$ consists of points which lie in the segment $\mathfrak{t}_k^3$ of an analytic hypersurface. Since further in $\mathfrak{A}^2$ the function $f = f[g_1(\zeta), g_2(\zeta)]$ is an analytic function of a complex variable ζ, the same inequality (5) holds in every interior point of the segment $\mathfrak{A}^2$.

In the Theorem 3.1 we assumed that f was regular in the (closed) domain $\overline{\mathfrak{M}}$. If we make some additional hypotheses about the structure of $\mathfrak{M}$, then it is possible to obtain an analogous result for functions which are regular in $\mathfrak{M}$ and continuous in $\overline{\mathfrak{M}}$.

We assume that for every t, $0 < t \leq p$, p sufficiently small

$$[z_\varkappa = h_\varkappa^k(Z_k, \lambda_k - it),\ Z_k \in \mathfrak{B}_k^2(\lambda_k, t),\ \varkappa = 1, 2], \qquad k = 1, 2, \ldots; n,$$

form the boundary of a domain $\mathfrak{M}_t$ which lies completely inside of $\mathfrak{M}$ and has the same structure as $\mathfrak{M}$. In particular, the $\mathfrak{B}^2(\lambda_k, t)$ are simply connected domains (which for $t = 0$ become unit circles). The boundary curves of the $\mathfrak{B}^2(\lambda_k, t)$, when t varies, change continuously.

For $t=0$, $[z_\varkappa = h_\varkappa^k(Z_k, \lambda_k), \varkappa = 1, 2, |Z_k|<1]$ are as before the segments of analytical hypersurfaces which form the boundary of $\mathfrak{M}$.

COROLLARY. If $f(z_1, z_2)$ is regular in $\mathfrak{M}$, continuous in $\overline{\mathfrak{M}}$, then in $\mathfrak{Z}_k^2(\lambda_k)$, $f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$ for every fixed λ_k is an analytic function of one complex variable Z_k, which is regular in $|Z_k|<1$.

Indeed, for $t>0$, the segment $z_\varkappa = h_\varkappa^k(Z_k, \lambda_k - i\,t)$ lies in $\mathfrak{M}$ and therefore $f[h_1^k(Z_k, \lambda_k - i\,t), h_2^k(Z_k, \lambda_k - i\,t)]$ is an analytic function of Z_k. Therefore for every closed curve $\mathfrak{C}^1$ in the Z_k plane we have

$$\int_{\mathfrak{C}^1} f[h_1^k(Z_k, \lambda_k - i\,t), h_2^k(Z_k, \lambda_k - i\,t)]\, dZ_k = 0.$$

Since f is assumed to be (uniformly) continuous the above integral must vanish also for $t=0$, and therefore by the MORRERA theorem $f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$ is an analytic function of the variable Z_k.

§ 2. A generalization of results of § 1 to the case of surfaces which intersect $\mathfrak{m}^3$ along a curve belonging to two segments, $\mathfrak{i}_k^3$ and $\mathfrak{i}_s^3$, $k \neq s$. We proceed now to the case where the segment $\overline{\mathfrak{A}^2}$ of an analytic surface intersects $\mathfrak{m}^3$ along curves which lie in two (or more) segments of analytic hypersurfaces $\mathfrak{i}_k^3$ and the intersection of their closures[7]), say in $\mathfrak{i}_1^3, \mathfrak{i}_2^3$, and $\mathfrak{i}_1^3 \cap \mathfrak{i}_2^3$.

More specifically, we shall assume that the intersection of an analytic surface $\mathfrak{A}_0^2$ with $\mathfrak{M}$ satisfies the following conditions.

1. *This intersection is a segment*

$$\overline{\mathfrak{A}^2} \equiv \mathfrak{A}_0^2 \cap \overline{\mathfrak{M}} = [z_\nu = g_\nu(\zeta),\ \nu = 1, 2,\ |\zeta| \leq 1]. \tag{3.6}$$

Here $g_\nu(\zeta)$ are analytic functions which are regular in $|\zeta|<1$ and continuous in $|\zeta| \leq 1$.

2. *The intersection of $\overline{\mathfrak{A}^2}$ with the boundary $\mathfrak{m}^3$ of $\mathfrak{M}$ consists of two segments $\mathfrak{a}_1^1$ and $\mathfrak{a}_2^1$, $\mathfrak{a}_k^1 < \overline{\mathfrak{i}_k^3}$, $k=1, 2$. The boundary curve $\mathfrak{a}^1$ of $\mathfrak{A}^2$ is $\mathfrak{m}^3 \cap \overline{\mathfrak{A}^2} = \sum_{k=1}^{2} \mathfrak{a}_k^1$. The representation of $\overline{\mathfrak{A}^2}$ in the ζ-plane implies that*

$$\mathfrak{a}_k^1 = [z_\nu = g_\nu(\exp(i\varphi)),\ \nu = 1, 2,\ \varphi_k \leq \varphi \leq \varphi_{k+1}] \tag{3.7}$$

where $k=1, 2$ and $\varphi_3 = \varphi_1 + 2\pi$.

3. *Since every point of $\mathfrak{a}_k^1$ lies also in a certain lamina, say $\mathfrak{Z}_k^2(\lambda_k) = [z_\nu = h_\nu(Z_k, \lambda_k), \nu = 1, 2]$ there exist continuous monotonous functions $\lambda_k = \lambda_k^*(\varphi)$ and $Z_k = Z_k^*[\lambda_k^*(\varphi)]$, $\varphi_k \leq \varphi \leq \varphi_{k+1}$ such that*

$$\mathfrak{a}_k^1 = [z_\nu = h_\nu^k[Z_k^*(\lambda_k^*(\varphi)), \lambda_k^*(\varphi)],\ \nu = 1, 2,\ \varphi_k \leq \varphi \leq \varphi_{k+1}]. \tag{3.8}$$

[7]) We assume here that $\mathfrak{A}^2$ intersects only two segments $\overline{\mathfrak{i}_k^3}$, and we make further a number of assumptions which simplify the formulation of the results. Most of these simplifications are unessential, and it is clear how they have to be modified in more general cases.

We assume that $\lambda_k^*(\varphi)$ *and* $Z_k^*(\lambda_k^*(\varphi))$ *are chosen in such a way that*

$$\lambda_k^*(\varphi^{(1)}) \neq \lambda_k^*(\varphi^{(2)}),\quad Z_k^*(\lambda_k^*(\varphi^{(1)})) \neq Z_k^*(\lambda_k^*(\varphi^{(2)})) \quad \text{for} \quad \varphi^{(1)} \neq \varphi^{(2)}.$$

THEOREM 3.2. *Let* f *be an analytic function of two complex variables, regular in* $\overline{\mathfrak{R}}$. *Let the intersection* $\overline{\mathfrak{A}^2} = [z_\nu = g_\nu(\zeta),\ \nu = 1, 2,\ |\zeta| \leq 1]$ *of* $\mathfrak{A}_0^2$ *with* $\overline{\mathfrak{R}}$ *satisfy the hypotheses 1–3. If, in addition, we assume that:*

4. *The function* $f(z_1, z_2)$ *omits the values* 0 *and* 1 *in the parts*

$$\sum_{\lambda_k = \lambda_k^*(\varphi_k)}^{\lambda_k^*(\varphi_{k+1})} \mathfrak{J}_k^2(\lambda_k) \quad \text{of} \quad \mathfrak{i}_k^3,\quad k = 1, 2.$$

5. *Along the curves* $\mathfrak{k}_k^1 = [z_\nu = h_\nu^k(0, \lambda_k),\ \nu = 1, 2]$, *where* $\lambda_k^*(\varphi_k) \leq \lambda_k \leq \lambda_k^*(\varphi_{k+1})$, *the functions* $f[h_1^k(0, \lambda_k), h_2^k(0, \lambda_k)]$, $k = 1, 2$, *assume values belonging to a bounded (closed) domain* $\overline{\mathfrak{P}_0^2}$ *that does not include the points* (0) *and* (1).

6. *The domain* $\mathfrak{P}_0^2$ *is such that for the functional* $A(\mathfrak{P}_0^2)$ *described in Theorem 3.1 and for* $|\varphi - \varphi_k| < \alpha$, $k = 1, 2$, *the inequality*

$$\left[\exp\left(\frac{A(\mathfrak{P}_0^2)}{1 - |Z_k^*(\lambda_k^*(\varphi))|}\right)\right] < C\,|\varphi - \varphi_k|^{-a}, \qquad 0 < a < 1, \tag{3.9}$$

holds. Here α *and* C *are conveniently chosen positive constants.*

Then at every point of $\mathfrak{A}^2$, *say at* $[z_\nu^0 = g_\nu(\zeta_0),\ \nu = 1, 2]$ *the function* f *satisfies the inequality*

$$|f(z_1^0, z_2^0)| \leq \left\{\exp\left[\frac{A(\mathfrak{P}_0^2)}{1 - r} + \frac{2C\alpha^{1-a}}{\pi(1-a)}\right]\right\}\left(\frac{1 + |\zeta_0|}{1 - |\zeta_0|}\right) \tag{3.10}$$

where r *is a conveniently chosen positive constant* $(r < 1)$.

PROOF. Our proof proceeds in a similar way as in the case of theorem 3.1.

We divide the segment $\mathfrak{a}^1$ into two parts[8]): $\sum_{k=1}^{2} \mathfrak{s}_k^1$ where

$$\mathfrak{s}_k^1 = [z_\nu = g_\nu^{(k)}(\exp(i\varphi)),\ \nu = 1, 2,\ \varphi_k - \alpha < \varphi < \varphi_k + \alpha] \tag{3.11}$$

and the complementary parts of $\mathfrak{a}^1$, which we shall denote by $\mathfrak{b}^1$.

Along $\mathfrak{s}_k^1$, the inequality (3.9) holds. In every lamina $\mathfrak{J}_k^2(\lambda_k)$, $\lambda_k = \lambda_k^*(\varphi)$, $\varphi_k - \alpha \leq \varphi \leq \varphi_k + \alpha$, the function $f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$ is a function of one complex variable Z_k defined in $|Z_k| \leq 1$ and omits there the values 0 and 1. It satisfies at $Z_k = 0$ the assumption 5., hence by SCHOTTKY's theorem (see [14], p. 198) we have for $\varphi_k - \alpha \leq \varphi < \varphi_k + \alpha$

$$\begin{cases} |f[h_1^k(Z_k^*(\lambda_k^*(\varphi)), \lambda_k^*(\varphi)),\ h_2^k(Z_k^*(\lambda_k^*(\varphi)), \lambda_k^*(\varphi))]| \\ \qquad \leq \exp\left[\dfrac{A(\mathfrak{P}_0^2)}{1 - |Z_k^*(\lambda_k^*(\varphi))|}\right] \leq C\,|\varphi - \varphi_k|^{-a}. \end{cases} \tag{3.12}$$

Since the remaining part $\mathfrak{b}^1 = \mathfrak{a}^1 - \sum_{k=1}^{2} \mathfrak{s}_k^1$, consists of two arcs, each of which lies completely in $\mathfrak{i}_k^3$ ($k = 1$ *or* 2), we conclude that for the representation

[8]) Each of these parts is a sum of two disconnected segments.

$Z_k = Z_k^*(\lambda_k)$ of $\mathfrak{b}^1$ in the Z_k-plane there exist[9]) $r = r(\alpha)$, $r < 1$, such that

$$|1 - |Z_k^*(\lambda_k)|| \geq 1 - r. \tag{3.13}$$

Therefore, from assumptions 4. and 5. follows the inequality (3.5) for values of $f(z_1, z_2)$, $(z_1, z_2) \in \mathfrak{b}^1$.

We consider now the function $f(z_1, z_2)$ in the segment $\overline{\mathfrak{A}^2}$, i.e. $f[g_1(\zeta), g_2(\zeta)]$, $|\zeta| \leq 1$. Considered as a function of ζ it is an analytic function regular in $|\zeta| < 1$, continuous in $|\zeta| \leq 1$. For $\zeta = \exp(i\varphi)$ it has the following bounds:

$$|f[g_1(\exp(i\varphi)), g_2(\exp(i\varphi))]| \leq \begin{cases} C|\varphi - \varphi_k|^{-a}, & \text{for } |\varphi - \varphi_k| \leq \alpha, \ k = 1, 2 \\ \exp\left[\frac{A(\mathfrak{P}_0^2)}{1-r}\right] & \text{for } |\varphi - \varphi_k| > \alpha. \end{cases} \tag{3.14}$$

The POISSON formula (see e.g. [10] p. 70) yields for $\zeta_0 \in \mathfrak{A}^2$:

$$\begin{cases} |f(z_1^0, z_2^0)| \equiv |f[g_1(\zeta_0), g_2(\zeta_0)]| \\ \qquad \leq \frac{1}{2\pi} \int_{\varphi=0}^{2\pi} |f[g_1(\exp(i\varphi)), g_2(\exp(i\varphi))]| \left| \frac{\exp(i\varphi) + \zeta_0}{\exp(i\varphi) - \zeta_0} \right| d\varphi \\ \qquad \leq \left[\exp\left(\frac{A(\mathfrak{P}_0^2)}{1-r}\right)\right] \left[\frac{1 + |\zeta_0|}{1 - |\zeta_0|}\right] + \frac{4C}{2\pi} \int_{\Phi=0}^{\alpha} \frac{1 + |\zeta_0|}{1 - |\zeta_0|} \frac{d\Phi}{\Phi^a} \end{cases} \tag{3.15}$$

which yields (3.10).

REMARK. In some instances instead of (3.14) it is convenient to use the inequality

$$|f[g_1(\exp(i\varphi)), g_2(\exp(i\varphi))]| \leq [A_0 + 2]^{\frac{\Delta}{1-r}} \tag{3.16}$$

where Δ is an absolute constant, and

$$A_0 = \max |f[h_k^1(0, \lambda_k), h_k^2(0, \lambda_k)]| \tag{3.17}$$

for $\lambda_k^*(\varphi_k) \leq \lambda_k \leq \lambda_k^*(\varphi_{k+1})$. See [10] p. 224.

REMARK. In the theorems 3.1 and 3.2 bounds for $f(z_1, z_2)$ in *one* segment, $\mathfrak{A}^2$, are given. Using a (two parameter) family $\mathfrak{A}^2(s)$ of surfaces we can cover $\mathfrak{M}$ (or a part of $\mathfrak{M}$) and using theorems 3.1 and 3.2 we obtain bounds for f in $\mathfrak{M}$ (or in a part of $\mathfrak{M}$). Using *some* families of surfaces one obtains immediately various theorems mentioned in the Introduction. In many instances one can investigate how to choose these families to obtain the best results.

§ 3. Bounds for $|f|$ if on every lamina $\mathfrak{J}_k^2(\lambda_k)$ of $\mathfrak{i}_k^3$, f is a univalent function. It is of considerable interest that various other results on value distribution of functions of *one* complex variable can be used in a similar manner to obtain theorems in the theory of two complex variables. For instance if in every lamina $\mathfrak{J}_k^2(\lambda_k)$ the function $f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$ is univalent (schlicht) we obtain the following result.

[9]) For instance we can choose $r = |Z_k^*(\lambda_k)|$.

THEOREM 3.3. *Suppose:* 1°. *The intersection of an analytic surface* $\mathfrak{A}_0^2 = [z_\nu = g_\nu(\zeta),\ \nu = 1, 2]$ *with the boundary* $\mathfrak{m}^3$ *of* $\mathfrak{M}$ *is a closed curve* $\mathfrak{p}_k^1 = [z_\nu = g_\nu(\exp(i\varphi)),\ 0 \leq \varphi \leq 2\pi]$ *which lies in the segment*[10] $\mathfrak{l}_k^3 = [z_\nu = h_\nu^k(Z_k, \lambda_k)]$.

2°. $\mathfrak{p}_k^1$ *lies in a segment which when its representation is normalized with respect to the axis* $\mathfrak{l}_k^1 = [z_\nu = h_\nu^k(0, \lambda_k)]$, *lies in the circle* $|Z_k| \leq R < 1$.

3°. *An analytic function* $f(z_1, z_2)$, $f[g_1(0), g_2(0)] = a$, *is given which is regular in* $\mathfrak{M}$ *and possesses the property that* $f[h_1(Z_k, \lambda_k), h_2(Z_k, \lambda_k)]$ *is for every fixed* λ_k *a univalent function of* Z_k *in* $|Z_k| \leq 1$.

4°. *Along the line* $\mathfrak{l}_k^1 = [z_\nu = h_\nu^k(0, \lambda_k),\ \nu = 1, 2]$

$$|f[h_1(0, \lambda_k), h_2(0, \lambda_k)]| \leq \alpha, \qquad \left|\frac{\partial f(z_1, z_2)}{\partial z_\nu}\right|_{z_\nu = h_\nu^k(0, \lambda_k)} \leq \beta_\nu, \tag{3.18}$$

$$\left|\frac{\partial h_\nu^k(Z_k, \lambda_k)}{\partial Z_k}\right|_{Z_k = 0} \leq \gamma_\nu, \qquad \nu = 1, 2. \tag{3.19}$$

Then

$$|f[g_1(\zeta), g_2(\zeta)] - a| \leq \left[\left(\alpha + |a| + \left(\sum_{\nu=1}^{2} \beta_\nu \gamma_\nu\right)\right) R(1-R)^{-2}\right] |\zeta|. \tag{3.20}$$

PROOF. Since for $|Z_k| < 1$, $f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]$ is univalent according to KOEBE's theorem we have

$$\frac{|Z_k|}{(1+|Z_k|)^2} \leq \left|\frac{f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)] - f[h_1^k(0, \lambda_k), h_2^k(0, \lambda_k)]}{B(0, \lambda_k)}\right| \leq \frac{|Z_k|}{(1-|Z_k|)^2}, \tag{3.21}$$

$$\frac{1-|Z_k|}{(1+|Z_k|)^3} \leq \left|\frac{B(Z_k, \lambda_k)}{B(0, \lambda_k)}\right| \leq \frac{1+|Z_k|}{(1-|Z_k|)^3}, \tag{3.22}$$

where

$$B(Z_k, \lambda_k) \equiv \frac{\partial f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]}{\partial Z_k}. \tag{3.23}$$

According to our hypotheses the values Z_k of the intersection curve

$$\mathfrak{A}^2 \cap \mathfrak{l}_k^3 = [z_\nu = g_\nu(\exp(i\varphi)) = h_\nu^k(Z_k, \lambda_k),\ \nu = 1, 2,\ 0 \leq \lambda_k \leq 2\pi,\ 0 \leq \varphi \leq 2\pi]$$

satisfy $|Z_k| \leq R$. Therefore by (3.21) at every point of $\mathfrak{A}^2 \cap \mathfrak{l}_k^3$ we have

$$\begin{cases} |f[g_1(\exp(i\varphi)), g_2(\exp(i\varphi))]| \equiv |f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)]| \\ \leq |f[h_1^k(0, \lambda_k), h_2^k(0, \lambda_k)]| + \left(\sum_{\nu=1}^{2} \frac{\partial f(z_1, z_2)}{\partial z_\nu} \frac{\partial h_\nu^k(Z_k, \lambda_k)}{\partial Z_k}\bigg|_{Z_k=0}\right) \frac{|Z_k|}{(1-|Z_k|)^2} \\ \leq \alpha + \left(\sum_{\nu=1}^{2} \beta_\nu \gamma_\nu\right) R(1-R)^{-2}. \end{cases} \tag{3.24}$$

Since according to the hypotheses $f[g_1(0), g_2(0)] = a$, the inequality (3.20) follows by SCHWARZ lemma.

THEOREM 3.4. *Suppose that the function* $f(z_1, z_2)$ *and the intersection* $\mathfrak{A}^2 = [z_\nu = g_\nu(\zeta),\ \nu = 1, 2,\ |\zeta| \leq 1]$ *with* $\mathfrak{m}^3$ *satisfy the conditions* 1° *and* 3° *of the theorem* 3.3.

[10]) Here and in the remaining part of this section we have to assume that the relations indicated in the footnote [6]) p. 186 hold.

Further, f satisfies 2°, with the inequality $|Z_k| \leq R$ being replaced by

$$1 > R \geq |Z_k| \geq r > 0. \tag{3.25}$$

Finally we assume

4°. *Along the axis $\mathfrak{k}_k^1$ the function $f(z_1, z_2)$ vanishes, i.e,*

$$f[h_1^k(0, \lambda_k), h_2^k(0, \lambda_k)] = 0. \tag{3.26}$$

5°. *For the function $B(Z_k, \lambda_k)$, see (3.23), it holds*

$$|B(0, \lambda_k)| \geq \tau > 0, \qquad 0 \leq \lambda_k \leq 2\pi, \tag{3.27}$$

where τ is a conveniently chosen constant.

6°. *$f(z_1, z_2)$ assumes on $\mathfrak{A}^2 = [z_\nu = g_\nu(\zeta), \nu = 1, 2, |\zeta| \leq 1]$ the value a only at the points*[11] $\zeta_1 = 0, \zeta_2, \ldots, \zeta_S$, $|\zeta_s| < 1$, $s = 2, 3, \ldots, S$.

Then

$$|f[g_1(\zeta), g_2(\zeta)] - a| \geq \tau r (1 + R)^{-2} \prod_{s=1}^{S} \left| \frac{\zeta - \zeta_s}{1 - \bar{\zeta}_s \zeta} \right|. \tag{3.28}$$

PROOF. Since $|B(0, \lambda_k)| > 0$, using the left hand side of the inequality (3.21) we obtain for the values $(z_1, z_2) \in \mathfrak{A}_0^2 \cap \mathfrak{i}_0^3$

$$|f[h_1^k(Z_k, \lambda_k), h_2^k(Z_k, \lambda_k)| \geq B(0, \lambda_k) |Z_k| (1 + |Z_k|)^{-2} \geq \tau r (1 + R)^{-2}. \tag{3.29}$$

Since

$$\left[\prod_{s=1}^{S} (\zeta - \zeta_s)/(1 - \zeta \bar{\zeta}_s) \right] [f(g_1(\zeta), g_2(\zeta)) - a]^{-1} \tag{3.30}$$

is regular in the unit circle $|\zeta| \leq 1$, (3.28) follows.

When formulating the theorem 3.4 we assumed that f vanishes on the axis $\mathfrak{k}_k^1$. It is possible to replace this requirement by a weaker one.

COROLLARY. Suppose that the function $f(z_1, z_2)$ and the intersection of $\mathfrak{A}_0^2$ with $\mathfrak{m}^3$ satisfy the conditions 1°, 2°, 3°, 4° of the theorem 3.4 (see also the theorem 3.3). Further we assume that $\mathfrak{A}_0^2 \cap \mathfrak{i}_k^3$ can be subdivided into finitely many parts,

$$\mathfrak{A}_0^2 \cap \mathfrak{j}_k^3 = \sum_{\sigma=1}^{K} \mathfrak{j}_\sigma^1, \tag{3.31}$$

$$\mathfrak{j}_\sigma^1 = [z_\nu = h_\nu^k(0, \lambda_k), \ \lambda_k^\sigma \leq \lambda_k < \lambda_k^{\sigma+1}, \ \sigma = 1, 2, \ldots, K], \qquad \lambda_k^{K+1} = \lambda_k^1 + 2\pi,$$

such that

$$|f(z_1, z_2)| \leq a_\sigma, \qquad |B(0, \lambda_k)| \geq \tau_\sigma \quad \text{on} \quad \mathfrak{j}_\sigma^1 \tag{3.32}$$

[see (3.23)]. Hereby we assume that the curve

$$\mathfrak{A}_0^2 \cap \mathfrak{i}_k^3 \equiv [z_\nu = h_\nu^k(\tilde{Z}_k(\lambda_k), \lambda_k), \ \nu = 1, 2]$$

and the axis

$$\mathfrak{k}_k^1 = [z_\nu = h_\nu^k(0, \lambda_k), \ \nu = 1, 2] = [z_\nu = g_\nu(\exp(i\varphi)), \ \nu = 1, 2],$$

[11] Every zero is counted according to its multiplicity.

are located in such a way that

$$(3.33) \qquad R_\sigma \geq |\tilde{Z}_k(\lambda_k)| \geq r_\sigma \quad \text{for} \quad \lambda_k^\sigma \leq \lambda_k \leq \lambda_k^{\sigma+1}, \qquad \lambda_k^{K+1} = \lambda_k^1 + 2\pi,$$

and

$$(3.34) \qquad a_\sigma < \tau_\sigma r_\sigma (1 + R_\sigma)^{-2}.$$

If, in addition, the condition 6° of the theorem 3.4 is satisfied and

$$g_\nu(\exp(i\varphi_\sigma)) = h_\nu[\tilde{Z}_k(\lambda_k^\sigma), \lambda_k^\sigma], \quad \nu = 1, 2$$

then for $(z_1, z_2) \in \mathfrak{A}_0^2 \cap \mathfrak{R}$ we have

$$(3.35) \quad \begin{cases} |f(z_1, z_2)| \equiv |f[g_1(\zeta), g_2(\zeta)]| \\ \qquad \geq \prod_{s=1}^{S} \left| \frac{\zeta - \zeta_s}{1 - \zeta \bar{\zeta}_s} \right| \exp\left[\sum_{\sigma=1}^{K} \log[\tau_\sigma r_\sigma (1 + R_\sigma)^{-2} - a_\sigma] \, H(\xi, \eta, \varphi_\sigma, \varphi_{\sigma+1}) \right] \end{cases}$$

where H is the harmonic measure[12]) of the arc $(\varphi_\sigma, \varphi_{\sigma+1})$ and $\zeta = \xi + i\eta$.

PROOF. We proceed in the same way as before. Using the left hand side of (3.21) and (3.33) we obtain that for $\lambda_k^\sigma \leq \lambda_k \leq \lambda_k^{\sigma+1}$, i.e. for $\varphi_\sigma \leq \varphi \leq \varphi_{\sigma+1}$,

$$(3.36) \quad \begin{cases} |f[g_1(\exp(i\varphi)), g_2(\exp(i\varphi))]| \\ \qquad = |f[h_1^k(\tilde{Z}_k(\lambda_k), \lambda_k), h_2(\tilde{Z}_k(\lambda_k), \lambda_k)]| \geq \tau_\sigma r_\sigma (1 + R_\sigma)^{-2} - a_\sigma. \end{cases}$$

Therefore for $|\zeta| < 1$ it holds

$$\log |f(g_1(\zeta), g_2(\zeta))| - \log\left(\prod_{s=1}^{S} \left| \frac{\zeta - \zeta_s}{1 - \zeta \bar{\zeta}_s} \right| \right) \geq \sum_{\sigma=1}^{K} \log[\tau_\sigma r_\sigma (1 + R_\sigma)^{-2} - a_\sigma] \, H(\xi, \eta; \varphi_\sigma, \varphi_{\sigma+1}).$$

Since in 6° we assumed that in $|\zeta| \leq 1$, $f[g_1(\zeta), g_2(\zeta)]$ vanishes only at the points ζ_s, $s = 1, 2, \ldots, S$, (3.35) follows.

REMARK. It should be added that the remark on p. 190 refers also to the theorems 3.3 and 3.4.

Bibliography.

[1] BERGMAN, S.: Über die Entwicklung der harmonischen Funktionen der Ebene und des Raumes nach Orthogonalfunktionen. Math. Ann. **86**, 237—271 (1922). — [2] BERGMAN, S.: Über die Veranschaulichung der Kreiskörper und Bereiche mit ausgezeichneter Randfläche. Jber. dtsch. Math.-Ver. **42**, 238—252 (1932/3). — [3] BERGMAN, S.: Über den Wertvorrat einer Funktion von zwei komplexen Veränderlichen. Math. Z. **36**, 171—183 (1932). — [4] BERGMAN, S.: Über eine in gewissen Bereichen gültige Integraldarstellung der Funktionen zweier komplexer Variabler. Math. Z. **39**, 76—94, 605—608 (1934). — [5] BERGMAN, S.: Über eine Integraldarstellung von Funktionen zweier komplexer Veränderlicher. Rec. Math. (Mat. Sbornik), N. S. **1** (43), 851—862 (1936). — [6] BERGMAN, S.: Über das Verhalten der Funktionen von zwei komplexen Veränderlichen in Gebieten mit einer ausgezeichneten Randfläche. Mathematica **14**, 107—123 (1938). — [7] BERGMAN, S.: Sur les fonctions orthogonales de plusieurs

[12]) I.e., that harmonic function of ξ, η, defined in $\xi^2 + \eta^2 \leq 1$ which equals 1 on the arc $(\varphi_\sigma, \varphi_{\sigma+1})$ and vanishes on the remaining part of the boundary of the unit circle.

variables complexes avec les applications à la théorie des fonctions analytiques. Mém. Sci. Math. **106**, 1—63 (1947). — [*8*] BERGMAN, S.: Sur la fonction-noyau d'un domaine et ses applications dans la théorie des transformations pseudo-conformes. Mém. Sci. Math. **108**, 1—80 (1948). — [*9*] BERGMAN, S.: Kernel function and extended classes in the theory of functions of complex variables. Colloque sur les fonctions de plusieurs variables, Bruxelles 1953, pp. 135—157. — [*10*] BIEBERBACH, L.: Lehrbuch der Funktionentheorie, Bd. 2. Leipzig 1931. — [*11*] BOCHNER, S.: Über orthogonale Systeme analytischer Funktionen. Math. Z. **14**, 180—207 (1922). — [*12*] BREMERMANN, H.: Analytic continuation of the kernel function and of the invariant metric in several complex variables. Proceedings of the Conference on Complex Variables. Ann. Arbor. Michigan **1954**. — [*13*] DAVIS, P.: On the application of doubly orthogonal functions to a problem of approximation in two regions. Trans. Amer. Math. Soc. **72**, 475—491 (1952). — [*14*] LITTLEWOOD, J. E.: Lectures on the theory of functions. Oxford: University Press 1944. — [*14a*] MASCHLER, M.: Minimal domains and the kernel function. To appear in Pacific Journal of Mathematics. — [*15*] MESCHKOWSKI, H.: Beiträge zur Theorie der Orthonormalsysteme. Math. Ann. **127**, 107—129 (1954). — [*16*] NEHARI, Z.: An integral equation associated with a function-theoretical extremal problem. To appear in J. d'Analyse Mathématique. — [*17*] SOMMER, F.: Über die Integralformeln in der Funktionentheorie mehrerer komplexer Veränderlichen. Math. Ann. **125**, 172—182 (1952). — [*18*] SZEGÖ, G.: Über orthogonale Polynome, die zu einer gegebenen Kurve der komplexen Ebene gehören. Math. Z. **9**, 218—270 (1921). — [*19*] WEIL, A.: L'intégrale de Cauchy et les fonctions de plusieurs variables. Math. Ann. **111**, 178—182 (1935). — [*20*] WEIL, A.: Sur les séries de polynomes de deux variables complexes. C. R. Paris **194**, 1904 to 1905 (1932).

Stanford University, Stanford, Cal. (USA.).

(Eingegangen am 14. März 1955.)

ROTHE, E. H.
Math. Zeitschr. Bd. 63, S. 195—218 (1955)

Mapping degree in Banach spaces and spectral theory.

To the memory of I. SCHUR in gratitude.

By

ERICH H. ROTHE.

1. INTRODUCTION. The mapping degree in a real Banach space R was established by Leray and Schauder[1]) for mappings of a bounded domain D of R into R which are of the form

$$y = f(x) = x - F(x) \tag{1.1}$$

where $F(x)$ is completely continuous. This was done by assuming Brouwer's theory of the mapping degree in finite dimensional Euclidean spaces as known, and then using an approximation procedure based on the fact that to every $\varepsilon > 0$ there exists a map $S(x)$ of D into a finite dimensional subspace of R such that

$$\|F(x) - S(x)\| < \varepsilon. \tag{1.2}$$

In the present paper an attempt is made to develop the degree theory directly in the Banach space R. Here already the case that (1.1) is a linear map requires some consideration due to the fact that no orientation is defined in a Banach space[2]). Accordingly, the main part of this paper deals with the linear case. The problem is to assign to a linear non-singular map (1.1) an index $j = j(f)$ which is $+1$ or -1, is invariant under homotopy, and $+1$ if f is the identity map I. It is natural to consider $x - \mu F(x)$ for $0 \leq \mu \leq 1$, or with $\lambda = \mu^{-1}$

$$f_\lambda(x) = (\lambda I - F)(x) \tag{1.3}$$

and the spectrum $\sigma(F)$ of F. If the part of $\sigma(F)$ with $\Re(\lambda) \geq 1$ [3]) contains only a finite number of points and if it is possible to assign to each such point

[1]) [6]. The theory was extended to linear convex topological spaces by Leray [5], and Nagumo [9]. (Numbers in brackets refer to the bibliography at the end of the paper.) The present paper is restricted to Banach spaces. As usual the norm of an element x of such a space will be denoted by $\|x\|$.

[2]) In a finite dimensional space the case of a linear map (1.1) is quite trivial, the degree being $+1$ or -1 according to whether the map (1.1) preserves the orientation or not, i.e., whether the determinant of (1.1) is > 0 or < 0. However, it seems to be an unsolved problem whether in case of a linear completely continuous Banach space map $F(x)$ the approximation (1.2) can be effected by a linear S. Therefore in order to establish the degree theory for linear maps (1.1) in the Banach space by the approximation method one would (at least at the present state of our knowledge) have to use the degree theory for non linear maps in the finite dimensional space.

[3]) $\Re(\lambda)$ denotes the real part of the complex number λ.

a positive integer m_i such that

$$j = \sum_i (-1)^{m_i} \tag{1.4}$$

is invariant under homotopy[4]), then (1.4) will represent a suitable definition for the index j[5]). This program is carried through in section 2. In this setup the homotopy theorem becomes a theorem on spectral perturbation, and, in order to use fully the spectral perturbation theory, we deal in that section with linear maps of a complex Banach space X into itself. In section 3 the special situation is treated where X admits a "conjugation" and where, consequently, it is possible to define "real" maps. Certain simplification in the basic assumptions of section 2 and in the definition (2.4) of the index can be made in this case. Section 4 deals with homotopy properties of real linear maps of such a space X into itself. Section 5 treats the non linear case in an arbitrary complex Banach space X, but only "smooth" mappings (def. 5.1) are considered while the treatment of mappings with less restrictive assumptions is left to further investigation. The results of section 5 for complex spaces are in section 6 carried over to real Banach spaces R by imbedding them (def. 6.1) in a complex Banach space.

2. THE INDEX OF A LINEAR MAP IN A COMPLEX BANACH SPACE X. Let X be a complex Banach space and F a bounded linear map of X into itself which satisfies the following assumptions A, B, and C:

ASSUMPTION A. No point of $\Re(\lambda) = 1$ belongs to the spectrum $\sigma(F)$ of F.

ASSUMPTION B. The intersection of $\sigma(F)$ with $\Re(\lambda) > 1$ consists of isolated points.

Before formulating Assumption C we remark first that B together with the boundedness of F implies that the intersection $\sigma(F) \cap \{\Re(\lambda) > 1\}$ consists of a finite number of points $\lambda_1, \lambda_2, \ldots, \lambda_\nu$, or is empty. Then, for $i = 1, 2, \ldots, \nu$, let c_i be the circumference of a counter-clockwise oriented circle with center λ_i and such that λ_i is the only point of $\sigma(F)$ inside or on c_i. Let E_{λ_i} be the projection[6]) defined by

$$E_{\lambda_i} = \frac{1}{2\pi\sqrt{-1}} \int_{c_i} R_\mu(F)\, d\mu \tag{2.1}$$

where

$$R_\mu(F) = (\mu I - F)^{-1}. \tag{2.2}$$

Finally, let

$$X_{\lambda_i} = E_{\lambda_i} X \tag{2.3}$$

be the range of E_{λ_i}.

[4]) It will be seen that these two properties are certainly true if F is completely continuous.

[5]) For the Leray-Schauder index the validity of (1.4) is a theorem (cf. [6] p. 58) while it will be a definition in the present paper.

[6]) [1], p. 196.

ASSUMPTION C. The dimension of X_{λ_i} is a finite number m_i [7]).

DEFINITION 2.1. The index $j = j(f)$ of the mapping (1.1) is defined by

$$j = j(f) = (-1)^{\sum_1^r m_i}. \tag{2.4}$$

The object of the present section is the proof of the following two theorems:

THEOREM 2.1. *Let $F_0(x)$ satisfy the assumptions* A, B, *and* C *and let*

$$f_0(x) = x - F_0(x). \tag{2.5}$$

Then there exists an $\varepsilon > 0$ of the following property: if the linear bounded map (1.1) *satisfies*

$$\|f - f_0\| = \|F - F_0\| < \varepsilon \tag{2.6}$$

then F satisfies the same assumptions as F_0 and

$$j(f) = j(f_0). \tag{2.7}$$

THEOREM 2.2. *For $0 \leq t \leq 1$ let $f_t(x) = x - F_t(x)$ where $F_t(x)$ satisfies assumptions* A, B, *and* C, *and where in addition F_t is continuous in t* [8]). *Then $j(f_1) = j(f_0)$.*

Theorem 2.2 is an immediate consequence of theorem 2.1 as application of the Heine-Borel theorem to the interval $0 \leq t \leq 1$ shows. The proof of theorem 2.1 is based on the following lemmas 2.1 and 2.2.

LEMMA 2.1. *Let F_0 satisfy the assumptions of theorem* 2.1. *We denote by $\sigma^+(F_0)$ and $\sigma^-(F_0)$ those parts of the spectrum $\sigma(F_0)$ of F_0 which lie in $\Re(\lambda) > 1$ and $\Re(\lambda) < 1$ respectively such that (by assumption* A) *$\sigma(F_0) = \sigma^+(F_0) + \sigma^-(F_0)$. Let 0^- be an open set in $\Re(\lambda) < 1$ containing $\sigma^-(F_0)$* [9]), *and let, for $i = 1, 2, \ldots, r_c$, c_i be counter-clockwise oriented circles defined as follows: by assumption* B, *$\sigma^+(F_0)$ consists of a finite number of points $\lambda_1^0, \lambda_2^0, \ldots, \lambda_r^0$. c_i is then a circle with center λ_i^0 which lies in $\Re(\lambda) > 1$ and such that the closed disk bounded by c_i contains no other point of $\sigma(F_0)$. Then there exists a positive ε_0 such that for*

$$\|F - F_0\| < \varepsilon_0 \tag{2.8}$$

the spectrum $\sigma(F)$ of F is contained in the open set 0 which is the union of 0^- and the open disks bounded by the c_i.

PROOF. Since 0 is an open set containing $\sigma(F_0)$ the lemma follows from the well-known fact that the spectrum $\sigma(F)$ of F depends continuously on F [10]).

LEMMA 2.2. *If F_0 is bounded linear and λ_i an isolated point of $\sigma(F_0)$ then there exists a positive ε_i such that for all linear bounded F satisfying*

$$\|F - F_0\| < \varepsilon_i \tag{2.9}$$

[7]) Assumptions A, B, and C are automatically satisfied if $F(x)$ is completely continuous. See, e.g., [1] p. 207.

[8]) I.e., for each t_0 in $0 \leq t \leq 1$ there exists to each $\eta > 0$ a ζ such that $\|F_t - F_{t_0}\| < \eta$ or what is the same $\|F_t(x) - F_{t_0}(x)\| < \|x\| \eta$ for $|t - t_0| < \zeta$.

[9]) Such set 0^- exists since $\sigma^-(F_0)$ is closed.

[10]) For definition and proof of the continuity of $\sigma(F)$, see [4] p. 118.

the following is true: if $m_i(F)$ denotes the dimension of the range of the projection

$$\frac{1}{2\pi\sqrt{-1}}\int_{c_i} R_\mu(F)\,d\mu \tag{2.10}$$

and if $m_i(F_0)$ is finite then $m_i(F)$ is finite and

$$m_i(F) = m_i(F_0). \tag{2.11}$$

We postpone the proof and show first that lemmas 2.1 and 2.2 together imply theorem 1. Using the notations used in the formulation of the two lemmas, we choose

$$\varepsilon = \min(\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r) \tag{2.12}$$

and will show that with this choice of ε, theorem 2.1 is true. That F satisfies assumption A is obvious from lemma 2.1. To prove assumption B we note first that lemma 2.1 also implies that the part of $\sigma(F)$ with $\Re(\lambda) > 1$ is contained in the open disks o_i bounded by the circles c_i. Since there are only a finite number of these disks it will be sufficient to prove that for fixed i the intersection $\sigma_i = o_i \cap \sigma(F)$ is finite. Now by the assumption made on F_0, $m_i(F_0)$ is finite (Assumption C). Therefore, by lemma 2.2, $m_i(F)$ is also finite; i.e., the range X_i of the projection (2.10) is a finite dimensional space. Now σ_i is a spectral set of F[11]). Therefore, by a theorem of Dunford[12]) σ_i is identical with the spectrum of the restriction F_{X_i} of F to the finite dimensional space X_i and therefore certainly finite. Thus assumption B is proved, and assumption C follows now easily: denote by $\lambda_1^i, \lambda_2^i, \ldots, \lambda_{i_n}^i$ the points of the set σ_i which just has been proved to be finite. Let o_j^i be disjoint circular disks with centers λ_j^i and lying in o_i. Let c_j^i be the counter-clockwise oriented boundary of o_j^i. Then

$$\frac{1}{2\pi\sqrt{-1}}\int_{c_i} R_\mu(F)\,d\mu = \sum_{j=1}^{i_n} \frac{1}{2\pi\sqrt{-1}}\int_{c_j^i} R_\mu(F)\,d\mu. \tag{2.13}$$

Therefore, X_i is the direct sum of the ranges X_j^i of the projections which are the terms of the right hand sum in (2.13) and

$$m_i(F) = \sum_{j=1}^{i_n} m_j^i(F) \tag{2.14}$$

if $m_j^i(F)$ denotes the dimension of X_j^i. Since $m_i(F)$ has already been proved to be finite, (2.14) implies that each $m_j^i(F)$ is finite which is assumption C. At the same time (2.7) follows now from (2.14), (2.11), and definition 2.1.

It remains to prove lemma 2.2. This lemma will be a consequence of four further lemmas.

---

[11]) A spectral set of F is a subset of $\sigma(F)$ which is open and closed in $\sigma(F)$. Cf. [1] p. 196.

[12]) [1] p. 197.

LEMMA 2.3. *Let B_0 be a linear bounded operator on X whose bounded inverse B_0^{-1} exists. Let B be a bounded linear operator which satisfies the inequality*

$$\|B - B_0\| < \tfrac{1}{2}\|B_0^{-1}\|^{-1}. \tag{2.15}$$

Then B^{-1} exists and satisfies the inequality

$$\|B^{-1} - B_0^{-1}\| < 2\|B_0^{-1}\|^2\|B - B_0\|. \tag{2.16}$$

PROOF. The fact that (2.15) implies the existence of B^{-1} and the inequality

$$\|B^{-1} - B_0^{-1}\| < \frac{\|B_0^{-1}\|^2\|B - B_0\|}{1 - \|B - B_0\|\|B_0^{-1}\|} \tag{2.17}$$

is well known[13]). But (2.17) together with (2.15) obviously implies (2.16).

LEMMA 2.4. *Let A_0 be a linear bounded operator on X. Let C be a closed bounded subset of the resolvent set $\varrho(A_0)$ of A_0. Then the resolvent $R_\lambda(A)$ of A is continuous at $A = A_0$ for $\lambda \subset C$, and the continuity is uniform as λ varies over C.*

PROOF. As is well known for each $x \subset X$, $R_\lambda(A_0)(x)$, and therefore also $\|R_\lambda(A_0)(x)\|$, is continuous for $\lambda \subset \varrho(A_0)$. Consequently, the latter real valued function of λ reaches a maximum in the closed bounded set C and we have for suitable positive M_x the inequality

$$\|R_\lambda(A_0)(x)\| < M_x \quad \text{for } \lambda \text{ in } C \text{ and all } x \subset X. \tag{2.16}$$

By a well-known theorem[14]) (2.16) implies the existence of a constant M such that

$$\|R_\lambda(A_0)\| < M \quad \text{for } \lambda \subset C. \tag{2.17}$$

Let now A be such that

$$\|A - A_0\| < \frac{1}{2M}. \tag{2.18}$$

We have then by (2.18) and (2.17)

$$\|(\lambda I - A_0) - (\lambda I - A)\| = \|A - A_0\| < \frac{1}{2M} < \frac{1}{2\|R_\lambda(A_0)\|} = \frac{1}{2\|(\lambda I - A_0)^{-1}\|}.$$

From this inequality we conclude by applying lemma 2.3 with $B_0 = \lambda I - A$ $B = \lambda I - A$ that $R_\lambda(A) = B^{-1}$ exists and satisfies the inequality

$$\|R_\lambda(A) - R_\lambda(A_0)\| < 2\|R_\lambda(A_0)\|^2\|B - B_0\| = 2\|R_\lambda(A_0)\|^2\|A - A_0\|$$

which because of (2.17) proves the lemma.

LEMMA 2.5. *Let c be a rectifiable curve in the resolvent set of A_0. Then*

$$\int_c R_\mu(A)\, d\mu$$

is continuous at $A = A_0$.

[13]) [4] p. 92.

[14]) See, e.g., [4] p. 26, Theorem 2.12.2.

PROOF. The lemma is an immediate consequence of the preceding lemma.

LEMMA 2.6[15]). *Let E_0 be a projection whose range has the finite dimension m_0, and let E be a projection for which*

$$\|E - E_0\| < \frac{1}{2\|E_0\|}. \tag{2.19}$$

Then the range of E has a finite dimension m and $m = m_0$.

PROOF OF LEMMA 2.2. Let E_0 be the projection obtained from (2.10) by replacing F by F_0. By lemma 2.5 we can choose an ε_i such that for F satisfying (2.9) the inequality (2.19) holds with E being the projection (2.10), and the equality (2.11) to be proved follows now from lemma 2.6.

3. REAL TRANSFORMATIONS IN A COMPLEX BANACH SPACE. In this section we consider "real" transformations in certain complex Banach spaces (definition 3.3)[16]). It turns out that for such transformations in the definition (2.4) of the index only those m_i have to be considered which correspond to real points of the spectrum (Theorem 3.1). In addition, some simplifications in assumptions A, B, and C of section 2 are possible (Theorem 3.2).

DEFINITION 3.1. Let X be a complex Banach space with elements $x, y, \ldots$. A map $x \to \bar{x}$ of X onto itself is called a conjugation if it satisfies the following four properties:

$$\bar{\bar{x}} = x \tag{3.1}$$

$$\overline{x + y} = \bar{x} + \bar{y} \tag{3.2}$$

$$\|x\| = \|\bar{x}\| \tag{3.3}$$

$$\overline{\lambda x} = \bar{\lambda}\bar{x} \tag{3.4}$$

where $\bar{\lambda}$ is the conjugate complex number to the complex number λ.

DEFINITION 3.2. Let $x \to \bar{x}$ be a conjugation in X. The elements x for which $x = \bar{x}$ are called real (with respect to the given conjugation).

LEMMA 3.1. *Let $x \to \bar{x}$ be a conjugation in X. Then the real elements of X (with their original norms) form a Banach space R over the reals.*

The proof consists in a rather obvious verification. We omit the details.

LEMMA 3.2. *Let $x \to \bar{x}$ be a conjugation in X, and let R be the real Banach space of the preceding lemma. Then every element x of X can be written uniquely as*

$$x = x' + i\,x'' \tag{3.5}$$

where x' and x'' are elements of R and $i = \sqrt{-1}$.

PROOF. It is easily seen that $x' = (x + \bar{x})/2$, $x'' = (x - \bar{x})/2i$ is the unique couple satisfying the requirements.

DEFINITION 3.3. Let X and R have the same meaning as in lemma 3.1. Then a continuous linear map A of X into itself or into another space with a conjugation is called real if

$$A(\bar{x}) = \overline{A(x)}. \tag{3.6}$$

[15]) For this lemma and its proof, see [10] p. 424, Lemma 4.

[16]) Cf. the similar definitions for the Hilbert space case given in [11], chapter IX, § 2.

LEMMA 3.3. *The continuous linear map A of X into itself is real if and only if it leaves R invariant.*

We omit the obvious proof.

LEMMA 3.4. *If A is real then* (i) *the spectrum $\sigma(A)$ of A is symmetric with respect to the real axis of the complex λ-plane and* (ii)

$$R_{\bar{\lambda}}(A)(x) = \overline{R_{\lambda}(A)(\bar{x})}. \tag{3.7}$$

PROOF. Since $\sigma(A)$ and the resolvent set $\varrho(A)$ of A are disjoint sets whose union is the whole λ-plane it will for the proof of i) be sufficient to show that $\varrho(A)$ is symmetric with respect to the real axis. Let then $\lambda_0 \subset \varrho(A)$; we have to prove that $\bar{\lambda}_0 \subset \varrho(A)$. By definition of the resolvent we have

$$(\lambda_0 I - A) R_{\lambda_0}(A)(x) = x \quad \text{for all} \quad x \subset X. \tag{3.8}$$

Since A is real it follows from lemma 3.3 that for all $y \subset X$

$$(\bar{\lambda}_0 I - A)(\bar{y}) = \overline{(\lambda_0 I - A)(y)}$$

and therefore from (3.8)

$$(\bar{\lambda}_0 I - A)\,\overline{R_{\lambda_0}(A)(x)} = \overline{(\lambda_0 I - A) R_{\lambda_0}(A)(x)} = \bar{x},$$

or replacing x by $\bar{x}$

$$(\bar{\lambda}_0 I - A)\,\overline{R_{\lambda_0}(A)(\bar{x})} = x. \tag{3.9}$$

We set now

$$S(x) = \overline{R_{\lambda_0}(A)(\bar{x})}. \tag{3.10}$$

$S(x)$ is bounded and as (3.9) shows a right inverse of $\bar{\lambda}_0 I - A$. Interchanging the two factors of the left member of (3.8) we would have seen that $S(x)$ is also a bounded left inverse. Thus $\bar{\lambda}_0 I - A$ has a uniquely determined bounded inverse which shows that $\bar{\lambda}_0$ is in $\varrho(A)$ and thus proves i). Moreover, since $S(x)$ is the bounded inverse of $\bar{\lambda}_0 I - A$ we see from the definition of the resolvent that $S(x) = R_{\bar{\lambda}_0}(A)(x)$. Comparison with (3.10) proves (3.7).

LEMMA 3.5. *Let A be real and λ_0 an isolated point of the spectrum $\sigma(A)$ of A such that by the preceding lemma $\bar{\lambda}_0$ is also an isolated point of $\sigma(A)$. Let c_0 be a counter-clockwise oriented circle with center λ_0 and such that the closed disk bounded by c_0 contains no other point of $\sigma(A)$ but λ_0, and let c_0' be the counter-clockwise oriented circle with center $\bar{\lambda}_0$ and of the same radius as c_0. Finally, let X_{λ_0}, $X_{\bar{\lambda}_0}$ denote the ranges of the projections*

$$E_{\lambda_0} = \frac{1}{2\pi\sqrt{-1}} \int_{c_0} R_{\mu}(A)\, d\mu, \qquad E_{\bar{\lambda}_0} = \frac{1}{2\pi\sqrt{-1}} \int_{c_0'} R_{\mu}(A)\, d\mu \tag{3.11}$$

respectively. We then claim:

i) *if λ_0 is not real then X_{λ_0} and $X_{\bar{\lambda}_0}$ are conjugate to each other (i.e., if y varies over X_{λ_0} then $\bar{y}$ varies over $X_{\bar{\lambda}_0}$), and, consequently, they are of the same dimension;*

ii) *if λ_0 is real then X_{λ_0} is invariant under conjugation.*

PROOF. Since $\bar{x}$ varies over all of X as x does, X_{λ_0} may be characterized as the set of all $E_{\lambda_0}(\bar{x})$ as x varies over X, and since $X_{\bar{\lambda}_0}$ is the set of all $E_{\bar{\lambda}_0}(x)$ as x varies over X, it is clear that for the proof of i) it is sufficient to prove

$$E_{\bar{\lambda}_0}(x) = \overline{E_{\lambda_0}(\bar{x})}. \tag{3.12}$$

But (3.12) implies also ii), for in the case of a real λ_0 it reads $\overline{E_{\lambda_0}(\bar{x})} = E_{\lambda_0}(x)$. Thus it remains to prove (3.12): let r be the common radius of c_0 and c_0'. We set $\mu = \bar{\lambda}_0 + r e^{i\varphi}, \lambda = \lambda_0 + r e^{-i\varphi}$ such that $\mu = \bar{\lambda}$. We then have from (3.11) and (3.7)

$$\begin{aligned} E_{\bar{\lambda}_0}(x) &= \frac{1}{2\pi}\int_0^{2\pi} R_\mu(A)(x)\, r e^{i\varphi}\, d\varphi = \frac{1}{2\pi}\int_0^{2\pi} R_{\bar{\lambda}}(A)(x)\, r e^{i\varphi}\, d\varphi \\ &= \frac{1}{2\pi}\int_0^{2\pi} \overline{R_\lambda(A)(\bar{x})\, r e^{-i\varphi}\, d\varphi} \\ &= -\frac{1}{2\pi i}\int_0^{2\pi} R_\lambda(A)(\bar{x})(-i r e^{-i\varphi}\, d\varphi) \\ &= -\frac{1}{2\pi i}\int_{-c_0} R_\lambda(A)(\bar{x})\, d\lambda = \overline{\frac{1}{2\pi i}\int_{c_0} R_\lambda(A)(\bar{x})\, d\lambda} = \overline{E_{\lambda_0}(\bar{x})}. \end{aligned}$$

LEMMA 3.6. *Let F satisfy the assumptions* A, B, *and* C *of section* 2 *and, in addition, be real. Let C be a simply closed rectifiable curve which is oriented in the counter-clockwise sense, lies in the half plane $\Re(\lambda) > 1$, and encloses in its interior the part of $\sigma(F)$ which lies in $\Re(\lambda) > 1$. Then the range of the projection*

$$E = \frac{1}{2\pi i}\int_C R_\mu(F)\, d\mu \tag{3.13}$$

is of finite dimension and invariant under conjugation.

PROOF. Let σ_C be the part of the spectrum $\sigma(F)$ which is enclosed by C. σ_C consists of a finite number of points. Now $\mu < \sigma_C$ implies $\bar{\mu} < \sigma_C$ by lemma 3.4. Therefore the points of σ_C can be written as $\mu_1, \bar{\mu}_1, \ldots, \mu_r, \bar{\mu}_r, \lambda_1, \ldots, \lambda_p$ where the μ's are the different non real points of σ_C and the λ's the different real points of σ_C. If then the projections $E_{\mu_j}, E_{\bar{\mu}_j}, E_{\lambda_j}$ are defined as in (2.1) we have

$$E = \sum_{j=1}^{r}(E_{\mu_j} + E_{\bar{\mu}_j}) + \sum_{j=1}^{p} E_{\lambda_j}. \tag{3.14}$$

It now follows from assumption C (section 2) that the range of E is of finite dimension and the invariance of this range is [because of (3.14)] a consequence of lemma 3.5.

THEOREM 3.1. *Let $f = I - F$ where F satisfies the assumptions of lemma* 3.6. *Let $\lambda_1, \ldots, \lambda_p$ be the real eigenvalues > 1 of F. Let E_{λ_i} be the projection* (2.1), *and m_i the dimension of its range. Then the index $j(f)$ of f is given by*

$$j(f) = (-1)^{\sum_{i=1}^{p} m_i} \tag{3.15}$$

PROOF. As in the proof of lemma 3.6 let $\mu_1, \bar{\mu}_1, \ldots, \mu_r, \bar{\mu}_r$ be the distinct non real eigenvalues of F with real part >1, and $E_{\mu_i}, E_{\bar{\mu}_i}$ the corresponding projections. Let m_i', m_i'' be the dimensions of the ranges of these projections. Then by definition (2.4)

$$j(f) = (-1)^s \tag{3.16}$$

where

$$s = \sum_{i=1}^{r} (m_i' + m_i'') + \sum_{i=1}^{p} m_i.$$

But by lemma 3.5, $m_i' = m_i''$. Therefore

$$s \equiv \sum_{i=1}^{p} m_i \mod 2,$$

and (3.15) follows from (3.16).

We will now show that for real transformations the set of assumptions A, B, C may be replaced by the following less restrictive set of assumptions:

Assumption A'. (1.1) is not singular [i.e., $\lambda = 1$ does not belong to $\sigma(F)$].

Assumption B'. $\sigma(F)$ is discrete for $\Re(\lambda) \geq 1$.

Assumption C' is the same as C except that the eigenvalues in question may have a real part $=1$ [17]).

THEOREM 3.2. *Theorems* 3.1, 2.1, *and* 2.2 *are still true if assumptions* A, B, C *are replaced by* A', B', C' *provided* F *is real.*

PROOF. Inspection of the proof of theorem 3.1 and the lemmas leading up to it shows that these proofs with obvious modifications still hold under the new assumptions if the curve C of lemma 3.6 is replaced by a simply closed rectifiable curve which encloses the part of $\sigma(F)$ with $\Re(\lambda) \geq 1$ and separates it from the part of $\sigma(F)$ with $\Re(\lambda) < 1$. It remains to prove theorem 2.1 since this theorem implies theorem 2.2. Let $f_0 = x - F_0$ where F_0 satisfies the assumptions of our theorem 3.2. If $\Re(\lambda) = 1$ contains no point of $\sigma(F_0)$ no new proof is required. Otherwise, $\sigma(F_0)$ contains a finite number of points on $\Re(\lambda) = 1$ which we denote by $\nu_1, \nu_2, \ldots, \nu_q$. As in the proof of lemma 3.5 let $\mu_1, \bar{\mu}_1, \ldots, \mu_r, \bar{\mu}_r$ be the distinct complex, and $\lambda_1, \ldots, \lambda_p$ the distinct real eigenvalues of F with $\Re(\lambda) > 1$ while $\sigma^-(F_0)$ denotes the set $\sigma(F_0) \cap \{\Re(\lambda) < 1\}$. Let then c_i, c_i', c_i'', c_i''' be circles with centers $\lambda_i, \mu_i, \bar{\mu}_i, \nu_i$ resp. of the following property: if o_i, o_i', o_i'', o_i''' are the open disks bounded by these circles then the closures of these disks are disjoint; for each of these closures its center is the only point belonging to $\sigma(F_0)$, and those of these closures whose centers are not on the real axis don't intersect the real axis. Let then 0^- be an open set in $\Re(\lambda) < 1$ containing $\sigma^-(F_0)$ and 0 be the union of 0^- and all the open disks defined above. Then there exists an ε such that

$$\|F - F_0\| < \varepsilon \tag{3.17}$$

implies

$$\sigma(F) < 0 \text{ [18])}. \tag{3.18}$$

[17]) However (because of assumption A') $\lambda = 1$ cannot be an eigenvalue.

[18]) Cf. footnote [10]).

By lemma 2.2 it is possible to choose ε in such a way that in addition to (3.18) the equality (2.11) holds with the $m_i(F)$ denoting the dimensions of the ranges of the projections (2.10) wher the c_i are the above circles with the real centers λ_i. Now by the choice of ε the intersection $\sigma(F) \cap \{\Re(\lambda) \geq 1\}$ is contained in the disks o_i, o_i', o_i'', o_i''', and since by construction only the o_i intersect the real axis, the real points of $\sigma(F) \cap \{\Re(\lambda) \geq 1\}$ are contained in the o_i. Since F is real the points of $\sigma(F) \cap \{\Re(\lambda) \geq 1\}$ contained in o_i may be written $\lambda_1^i, \ldots, \lambda_{p_i}^i, \mu_1^i, \overline{\mu_1^i}, \ldots, \mu_{r_i}^i, \overline{\mu_{r_i}^i}$ where the λ's are real and the μ's non real. To each of these points corresponds in the usual way a projection and we denote the dimensions of the ranges of these projections by $m_j^i(F)$, $n_j^i(F)$, $\bar{n}_j^i(F)$. Then an argument similar to the one leading to (2.14) shows that

$$m_i(F) = \sum_{j=1}^{p_i} m_j^i(F) + \sum_{j=1}^{r_i} (n_j^i(F) + \bar{n}_j^i(F)). \tag{3.19}$$

But $n_j^i(F) = \bar{n}_j^i(F)$ since F is real. Therefore we see from (3.19) and (2.11) that

$$m_i(F_0) \equiv \sum_{j=1}^{p_i} m_j^i(F) \mod 2.$$

By theorem 3.1 this equality implies (2.7).

REMARK 3.1. If $f(x) = x - F(x)$ is linear, non singular, real, and $F(x)$ completely continuous, then the assumptions A', B', C' follow. For the restriction of f to R (cf. lemma 3.1) an index $j_1(f)$ can then be defined by Leray-Schauder's method (cf. [6], section 11). It is easily seen that $j_1(f) = j(f)$ if one observes that the "degree m_i in the sense of Goursat" used in [6] is identical with the number m_i introduced in Assumption C of the present paper, and if one uses the "Conclusion" in [6], p. 58.

4. HOMOTOPY PROPERTIES. THE PRODUCT THEOREM.

DEFINITION 4.1. Let F, G be bounded linear real transformations mapping the complex Banach space X into itself and satisfying the assumptions A', B', C' of section 3. F and G are called homotopic if for each t in the interval $0 \leq t \leq 1$ there exists an F_t which satisfies all assumptions just mentioned for F and G, which is moreover continuous in t and such that $F_0 = F$, $F_1 = G$. The maps

$$f(x) = x - F(x), \quad g(x) = x - G(x) \tag{4.1}$$

are called homotopic if F and G are homotopic.

THEOREM 4.1. *Let the linear bounded maps F and G be real and satisfy the assumptions* A', B', C'. *Then the maps* (4.1) *are homotopic if and only if $j(f) = j(g)$. In particular: $j(f) = +1$ if and only if f is homotopic to the identity map, and $j(f) = -1$ if and only if f is homotopic to a map $g(x)$ of the form $g(x) = x - 2\varphi(x)\, x^0$ where x^0 is an arbitrarily chosen real element of X of norm 1, and $\varphi(x)$ an arbitrary linear real functional with $\varphi(x^0) = 1$.*

PROOF. That $j(f)=+1$ if f is homotopic to the identity, and $j(f)=-1$ if $f=g$ is a consequence of theorem 3.1 since the index of the identity map is $+1$, and $j(g)$ is easily seen to be -1. Before proceeding to prove the converse we need a few simple facts which we state as lemmas.

LEMMA 4.1. *Let X^0 be a finite dimensional subspace of X, $\tilde{F}^0$ a real linear map of X^0 into itself and I^0 the identity map of X^0. The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $\tilde{F}^0$ are supposed to be all real and >1. Then $f^0=I^0-\tilde{F}^0$ is homotopic to $-I^0$.*

PROOF. We choose the coordinate system in X^0 in such a way that the matrix M representing $\tilde{F}_0$ in this coordinate system is in the Jordan normal form, i.e., that M is direct sum of matrices each of which contains one of the λ_i in the main diagonal and (unless it is one dimensional) contains one's directly above the main diagonal while all its other elements are zero. We define for $0\leq t\leq 1$ as M_t the matrix obtained from M by replacing each λ_i by $\lambda_i(1-t)+2t$ and the one's by $1-t$, and by $\tilde{F}_t^0$ the transformation given by M_t. Then $f_1^0=I^0-\tilde{F}_1^0$ has the matrix $I^0-2I^0=-I^0$ and is homotopic to $f^0=I^0-\tilde{F}_0^0$.

LEMMA 4.2. *With the notations and assumptions of lemma 4.1 let N be the dimension of X^0. Then:* i) *if N is even, f^0 is homotopic to the identity I^0;* ii) *if N is odd, $\tilde{f}^0$ is homotopic to a map of the form $g(x)=x-2\varphi(x)\,x^0$ where x^0 is an arbitrarily chosen element of X^0 of norm 1 and $\varphi(x)$ a linear functional with $\varphi(x^0)=1$;* iii) *if X^0 is invariant under conjugation then x^0 and $\varphi(x)$ may be chosen to be real.*

PROOF. On account of lemma 4.1 we may assume that $\tilde{f}^0=-I^0$. Now it is well known that in a space of even dimension the negative identity map is homotopic to the identity map. Applying this fact to X^0 proves i), and applying it to the even dimensional space which together with x^0 spans X^0 proves obviously ii). Under the assumption made in iii), x^0 obviously can be chosen real[19]). We define $\varphi(x)$ first for $x=\alpha x^0$ by setting $\varphi(x)=\alpha$. Since x^0 is real $\varphi(x)$ is real. By the following lemma 4.3 we can extend $\varphi(x)$ to X as a real linear functional.

LEMMA 4.3. *Let X be a complex Banach space with a conjugation and X_1 a finite dimensional linear subspace which is invariant under this conjugation, and $l(x)$ a linear functional defined on X_1 which is real in the sense of definition 3.3. Then $l(x)$ can be extended to X with the original bound L as a real linear functional.*

PROOF. Let R be the real Banach space consisting of the real elements of X (cf. lemma 3.1). It is then easily seen that the intersection $R_1=R\cap X_1$ is a finite dimensional linear subspace of R. Therefore we can by the Hahn-Banach theorem for real Banach spaces extend $l(x)$ as a real valued linear functional to R with the original bound L. Let now x be an arbitrary element

[19]) Let $x\neq 0$ be an element of X^0. If then in the decomposition (3.5) $x'\neq 0$ we set $x^0=(x+\bar{x})/\|x+\bar{x}\|$. If $x'=0$ we replace x by ix.

of X. By lemma 3.2 we can write uniquely

$$x = x' + i\,x'' \qquad (x', x'' < R).$$

We then define

$$l(x) = l(x') + i\,l(x'').$$

It is a matter of simple computation to verify first the additivity of $l(x)$ and then the equality $l(\alpha x) = \alpha\, l(x)$ for arbitrary complex α. To prove the boundedness of $l(x)$ with the same bound L, let

$$l(x) = |l(x)|\, e^{-i\vartheta}, \qquad y = x\, e^{i\vartheta} = y' + i\,y'' \quad \text{with} \quad y', y'' < R.$$

Then

$$|l(x)| = l(x\, e^{i\vartheta}) = l(y) = l(y') + i\,l(y'').$$

Since $l(y'), l(y'')$ are real numbers and $y' < R$, this implies

$$|l(x)| = l(y') \leq L\|y'\| = L\|y + \bar{y}\|/2 \leq L\,\frac{\|y\| + \|\bar{y}\|}{2} = L\|y\| = L\|x\|.$$

Finally, it is immediately verified that $l(x)$ is real, i.e., satisfies $l(\bar{x}) = \overline{l(x)}$.

LEMMA 4.4. *Let X be the direct sum of the subspaces X^0 and X^1, and E^0 and E^1 the projections of X onto X^0 and X^1 resp. Let G^i be a bounded linear map of X^i into itself $(i = 0, 1)$. Define for $x < X$ the map $G(x) = G^0E^0(x) + G^1E^1(x)$. Then $\sigma(G)$ is the union of $\sigma(G^0)$ and $\sigma(G^1)$.*

PROOF. Since for any bounded transformation T the spectrum $\sigma(T)$ and the resolvent set $\varrho(T)$ are complimentary sets (with respect to the whole λ-plane) the statement of the lemma is equivalent to the equality

$$\varrho(G) = \varrho(G^0) \cap \varrho(G^1).$$

But this equality is easily derived from the fact the equation $\lambda x - G(x) = y$ is equivalent to the pair of equations for $x^0 = E^0 x$ and $x^1 = E^1(x)$

$$\lambda\, x^0 - G^0(x^0) = E^0(y), \qquad \lambda\, x^1 - G^1(x^1) = E^1(y)$$

together with the theorem[20]) that an everywhere defined inverse of an everywhere defined bounded linear transformation is itself bounded.

We now return to the proof of theorem 4.1. Let Γ^0 and Γ^1 be two simply closed rectifiable and disjoint curves oriented in the counter clockwise sense of the following properties: i) they lie in the resolvent set $\varrho(F)$ of F; ii) the open domains D^0 and D^1 bounded by them are disjoint; iii) let σ^0 be the part of $\sigma(F)$ which is real and > 1; then there exists an interval of the real axis which contains the point $\lambda = 2$ and σ^0, and which itself is contained in D^0; iv) $\sigma^1 = \sigma(F) - \sigma^0$ lies in D^1. We set

$$(4.2) \qquad F^0 = \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma^0} \mu\, R_\mu(F)\, d\mu, \qquad F^1 = \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma^1} \mu\, R_\mu(F)\, a\mu$$

$$(4.3) \qquad E^0 = \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma^0} R_\mu(F)\, d\mu, \qquad E^1 = \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma^1} R_\mu(F)\, d\mu$$

$$(4.4) \qquad X^0 = E^0 X, \qquad X^1 = E^1 X.$$

[20]) See e.g. [4], p. 29, Corollary.

Then, as well known[21]), F^0 leaves X^0, F^1 leaves X^1 invariant, the restriction $\tilde{F}^0$ of F^0 to X^0 has σ^0 as its spectrum, and X is the direct sum of X^0 and X^1. σ^0 consists by assumption of a finite number of points $\lambda_1, \ldots, \lambda_n$. If m_i is the dimension of the range of the projection (2.1) and N the dimension of X^0 comparison of (2.1) with (4.3) and (4.4) shows that $N = \sum_{i=1}^{n} m_i$. Consequently, we see from theorems 3.1, 3.2 that

$$j(f) = (-1)^N. \tag{4.5}$$

Now X^0 is invariant under conjugation by lemma 3.6 and $\tilde{F}^0$ satisfies the assumptions of lemma 4.2[22]). Therefore there exists for $0 \leq t \leq 1$ a map $\tilde{F}_t^0$ of X^0 into itself (satisfying A', B', C') such that $\tilde{F}_0^0 = \tilde{F}^0$ while

$$\tilde{f}_1^0(x) = x - \tilde{F}_1^0(x) = \begin{cases} x & \text{if } N \text{ is even} \\ x - 2\varphi(x)\, x^0 & \text{if } N \text{ is odd} \end{cases} \tag{4.6}$$

where x^0, φ are as in lemma 4.2. Let now $\tilde{F}^1$ be the map F^1 [defined in (4.2)] restricted to X^1, and $\tilde{F}_t^1 = (1-t)\tilde{F}^1$ for $0 \leq t \leq 1$. The spectrum $\sigma(\tilde{F}_t^1)$ of $\tilde{F}_t^1$ is obtained in multiplying the points of $\sigma(\tilde{F}^1)$ by $1-t$, and since the real points of $\sigma(\tilde{F}^1) = \sigma(F^1)$ are in the half plane $R(\lambda) < 1$ and $\tilde{F}^1$ satisfies A', B', C', it follows easily that also $\tilde{F}_t^1$ satisfies A', B', C'. We now define for arbitrary x in X

$$F_t(x) = \tilde{F}_t^0 E^0(x) + \tilde{F}_t^1 E^1(x).$$

Since by lemma 4.4, $\sigma(F_t)$ is the union of $\sigma(\tilde{F}_t^0)$ and $(\tilde{F}_t^1)$ it follows that F_t satisfies assumptions A', B', C', and $f(x) = x - F(x) = x - F_0(x)$ is by (4.6) seen to be homotopic to

$$\begin{aligned} f_1(x) = x - F_1(x) &= \left(E^0(x) - \tilde{F}_1^0(x)\right) + E^1(x) = \tilde{f}_1^0 E^0(x) + E^1(x) \\ &= x \qquad \text{if } N \text{ is even} \\ &= x - 2\varphi(x)\, x^0 \quad \text{if } N \text{ is odd.} \end{aligned}$$

It remains to prove that x^0 and $\varphi(x)$ can be chosen arbitrarily subject to the conditions of reality and to

$$\|x^0\| = 1, \qquad \varphi(x^0) = 1, \tag{4.7}$$

i.e., we have to show: if

$$f(x) = x - 2\varphi(x)\, x^0, \qquad g(x) = x - 2\psi(x)\, y^0$$

where x^0, y^0 are real elements of X and $\varphi(x)$, $\psi(x)$ real linear functionals satisfying (4.7) and

$$\|y^0\| = 1, \qquad \psi(y^0) = 1, \tag{4.8}$$

then f and g are homotopic. We distinguish two cases:

CASE 1. x^0 and y^0 are linearly dependent. In this case $y^0 = \varepsilon x^0$ where because of (4.7), (4.8), ε is a complex number of absolute value 1. Then

[21]) [1], p. 197.

[22]) That $\tilde{F}^0$ is real follows from the assumption that F is the real since $\tilde{F}^0$ is the restriction of F to X^0.

$\psi(x)\,y^0 = \tilde{\psi}(x)\,x^0$ if $\tilde{\psi}(x)$ is defined as $\psi(\varepsilon x)$, and $\tilde{\psi}(x^0) = \psi(\varepsilon x^0) = \psi(y^0) = 1$, which shows that we may assume that $\varepsilon = 1$, i.e., $y^0 = x^0$. With this assumption we set $\varphi_t(x) = (1-t)\,\varphi(x) + t\,\psi(x)$ for $0 \leq t \leq 1$. We note that $\varphi_t(x^0) = 1$ and that therefore the spectrum of $F_t(x) = 2\varphi_t(x)\,x^0$ consists of the points $\lambda = 0$ and $\lambda = 2$, and that the range of the projection belonging to $\lambda = 2$ is of dimension 1. Therefore $F_t(x)$ satisfies A', B', C'. Moreover, with $f_t(x) = x - F_t(x)$ we have $f_0 = f$, $f_1 = g$. Thus f and g are homotopic in the sense of definition 4.1.

CASE 2. x^0 and y^0 are linearly independent. Let X^2 be the subspace of X spanned by x^0 and y^0, and E^2 a real projection with range X^2 [23]). Let for $x \subset X^2$, $\tilde{r}_t(x)$ be a rotation about the origin such that $\tilde{r}_0(x^0) = x^0$, $\tilde{r}_1(x^0) = y^0$ [24]), let $r_t = \tilde{r}_t / \|\tilde{r}_t\|$, and let the extension of r_t to all of X be defined by

$$r_t(x) = x - E^2(x) + r_t\big(E^2(x)\big).$$

Moreover, set

$$\varphi_t(x) = \varphi\big(r_t^{-1}(x)\big), \qquad F_t(x) = 2\varphi_t(x)\,r_t(x^0).$$

If one observes that $\varphi_t\big(r_t(x^0)\big) = \varphi(x^0) = 1$, it is easily seen that the spectrum of F_t consists of the points $\lambda = 0$ and $\lambda = 2$, and that F_t satisfies A', B', C'. Therefore setting $f_t(x) = x - F_t(x)$ shows that $f(x)$ is homotopic to $h(x) = x - 2\varphi_1(x)\,y^0$ since $r_1(x^0) = y^0$. To see finally that h is homotopic to g we have only to notice that we are now in case 1 [25]).

LEMMA 4.5. *Let X be a complex Banach space with a conjugation, and X_p a p-dimensional linear subspace which is invariant under this conjugation (p a finite integer). Then there exists a real projection E_p of X whose range is X_p (cf. definition 3.3).*

PROOF. Without the reality requirement E_p may, according to Mackey [26]), be constructed as follows: let $x_1, x_2, \ldots, x_p$ be a base for X_p. Then for $j = 1, \ldots, p$ the coefficient $f_j(x)$ of x_j in the representation of an element x of X as linear combination of $x_1, \ldots, x_p$ is a linear functional on X_p which may be extended to X. Calling the extension again $f_j(x)$, it turns out that $E_p(x) = \sum_{j=1}^{p} x_j\,f_j(x)$ is a projection with range X_p. We will now show that with a proper choice of the basis $x_1, \ldots, x_p$, E_p is real. We will prove first that it is possible to construct a basis of real elements: let $x \neq 0$ be an element of X_p; then by assumption $\bar{x} \subset X_p$, and therefore $x + \bar{x}$ and $(x - \bar{x})/i$ are real element of X_p at least one of them not being the zero element. Thus if q denotes the maximal number of real elements of X_p which are linearly independent (with respect to complex numbers) then $q \geq 1$. We will prove $q = p$: let $r_1, r_2, \ldots, r_q$, be linearly independent real elements of X_p, and

[23]) Such E^2 exists by the following lemma 4.5.

[24]) i.e., if $x = \alpha^0 x^0 + \alpha' y^0$ then $\tilde{r}_t(x) = \beta^0 x^0 + \beta' y^0$ where $\beta^0 = \alpha^0 \cos(t\pi/2) - \alpha' \sin(t\pi/2)$, $\beta' = \alpha^0 \sin(t/\pi/2) + \alpha' \cos(t\pi/2)$.

[25]) Note that $\varphi_1(y^0) = \varphi_1(r_1(x^0)) = \varphi(x^0) = 1$.

[26]) [7], lemma 3.1.

$x = x' + ix''$ (x', x'' real) an arbitrary element of X_p. It follows from these definitions that there exist constants α_j', α_j'' (which may be complex) such that

$$x' = \sum_{j=1}^{q} \alpha_j' r_j, \qquad x'' = \sum_{j=1}^{q} \alpha_j'' r_j.$$

Then $x = x' + ix''$ is a linear combinations of the r_j which proves that the r_j form a base and $q = p$. Thus we may assume that the elements $x_1, \ldots, x_q$ of the original base are all real. But then comparison of

$$\bar{x} = \sum_{j=1}^{q} f_j(\bar{x})\, x_j \quad \text{with} \quad \bar{x} = \overline{\sum_{j=1}^{q} f_j(x)\, x_j} = \sum_{j=1}^{q} \overline{f_j(x)}\, x_j$$

shows that $f_j(\bar{x}) = \overline{f_j(x)}$, i.e., that the functionals $f_j(x)$ are real. Therefore, by lemma 4.3 their extensions can also be constructed to be real, and the above constructed projection E_p is then obviously also real.

REMARK 4.2. It is well known[27]) that the range X_{λ_i} of the projection (2.1) is the space spanned by the solutions of the totality of equations $(\lambda_i I - F)^n = 0$ for $n = 1, 2, \ldots$. If now F and λ_i is real then X_{λ_i} is invariant under conjugation (lemma 3.5). It follows (as shown in the proof of lemma 4.5) that X_{λ_i} can be spanned by elements of R, in other words, the dimension m_i of X_{λ_i} is the number of linearly independent solutions in R of the above equations. It follows then from theorem 3.1 that the index $j(f)$ is determined if F is known in R. If F is in addition completely continuous it also follows that $j(f)$ agrees with the Leray-Schauder index of f (cf. [6] p. 57, 58).

THEOREM 4.2. (*Product theorem.*) *Let* $f(x) = x - F(x)$ *and* $g(x) = x - G(x)$. *Let* $F(x)$, $G(x)$, *and* $x - (fg)(x)$ *be real and satisfy* A', B', C'. *Then*

$$j(fg) = j(f)\, j(g). \tag{4.9}$$

PROOF. If one of the mappings, say f, has index $+1$, then by theorem 4.1 f is homotopic to the identity map and therefore fg is homotopic to g. We have then, again using theorem 4.1, $j(fg) = j(g) = 1 \cdot j(g) = j(f)\, j(g)$, which proves (4.9) in the case considered. Assume now

$$j(f) = j(g) = -1. \tag{4.10}$$

Then by theorem 4.1, f and g are both homotopic to a map of the form

$$h(x) = x - 2\varphi(x)\, x^0 \quad \text{with} \quad \|x^0\| = \varphi(x^0) = 1. \tag{4.11}$$

Then fg is homotopic to h^2, and by theorem 4.1, we have $j(fg) = j(h^2)$. This equality together with (4.10) shows that for the proof of (4.9) it will be sufficient to show that

$$j(h^2) = +1. \tag{4.12}$$

But (4.12) is certainly true since $h^2(x) = x$ as is easily verified from (4.11).

[27]) See e.g., [1], p. 198.

5. NON LINEAR MAPS. Let X be a complex Banach space, D a bounded open set in X, $\overline{D}$ its closure, and $B=\overline{D}-D$ its boundary. We will define the mapping degree and prove the homotopy property for a class of "smooth" mappings of $\overline{D}$ into X.

DEFINITION 5.1. A mapping

$$f(x)=x-F(x) \tag{5.1}$$

of $\overline{D}$ into X is called smooth if it satisfies the following conditions: i) in each point x of $\overline{D}$, $f(x)$ has a bounded Fréchet differential $d(x,h)$[28]) which is uniformly continuous in x in the sense that to each positive ε there exists a δ such that

$$\|d(x,h)-d(x_0,h\|<\varepsilon\|h\| \quad \text{for} \quad \|x-x_0\|<\delta; \tag{5.2}$$

ii) $d(x,h)$ (as linear operator in h) has an inverse $d^{-1}(x,h)=l(x,h)$ with a bound which is uniform for all $x\subset\overline{D}$; iii) if the point y of X is not on the image of the boundary B of D, then there are at most a finite number of points x in D which satisfy the equation

$$f(x)=y; \tag{5.3}$$

iv) the linear operator in h, $D(x,h)=h-d(x,h)$ satisfies the conditions A, B, C of section 2; v) the image of a closed set is closed.

REMARK 5.1. If only i) and ii) are assumed and $F(x)$ is completely continuous then v) is automatically satisfied; so is iii), for it can be easily seen that otherwise there would be a point x in D satisfying (5.3) in every neighborhood of which there would be other solutions of (5.3) which would be a contradiction to the lemma 5.1 below. If in addition $D(x,h)$ is completely continuous (in h) then iv) is also automatically satisfied.

DEFINITION 5.2. Let $f(x)$ be a smooth mapping of $\overline{D}$ into X, and y a point of X which is not on the image of the boundary B of D. We define the degree $j(f,y,D)$ of f at y with respect to D as follows: if (5.3) has no solution x in D we set $j(f,y,D)=0$; if there are solutions in D there are by assumption iii) of definition 5.1 only a finite number, say $x_1, x_2, \ldots, x_r$. By assumption iv) of definition 5.1 for each of the linear operators in h, $d(x_\varrho,h)$ $(\varrho=1,\ldots,r)$ the index $j_\varrho=j\big(d(x_\varrho,h)\big)$ is defined according to definition 2.1. Let p be the number of those j_ϱ which are $+1$, and q the number of those which are -1. Then we define

$$j(f,y,D)=p-q \text{ [29]}. \tag{5.4}$$

[28]) For fixed x, $d(x,h)$ is a bounded linear operator in h.

[29]) In the Leray Schauder theory the equality (5.4) is a theorem (cf. [6], p. 56). In finite dimensional spaces p is the numbers of those x for which the Jacobian of $f(x)$ is positive, and q the number of those for which it is negative. With this interpretation (5.4) is the basic definition used as starting point in Nagumo's theory for the degree in finite dimensional spaces. (See [8].)

DEFINITION 5.3. Let $f(x)=x-F(x)$, $g(x)=x-G(x)$ be two smooth mappings of $\overline{D}$ into X. Then f and g are called homotopic if for each t in the interval $0\leq t\leq 1$ there exists a smooth map

$$f_t(x)=x-F(x,t) \quad \text{with} \quad f_0(x)=f(x), \quad f_1(x)=g(x) \tag{5.5}$$

where $F(x,t)$ is continuous in the product $\Pi=\overline{D}\times\{0\leq t\leq 1\}$, where the continuity in t for fixed x is uniform for $x\subset\overline{D}$, and where moreover the following two conditions are satisfied: a) the Fréchet differential $d(x,t,h)$ of $f_t(x)$ (with respect to x) is continuous in x in the sense that to every positive ε there exist positive numbers δ and τ such that

$$\|d(x,t,h)-d(x_0,t_0,h\|<\varepsilon\|h\| \quad \text{for} \quad \|x-x_0\|<\delta, \quad |t-t_0|<\tau;$$

b) the inverse $d^{-1}(x,t,h)=l(x,t,h)$ of $d(x,t,h)$ as linear operator in h exists and is bounded uniformly on $\overline{D}$, i.e., there exists a positive m such that

$$\|l(x,t,h)\|<m\|h\|. \tag{5.6}$$

THEOREM 5.1. *Let f, g be two smooth mappings of $\overline{D}$ into X, and y a point of X which does not lie on the image of the boundary B of D under f or g. We assume that f and g are homotopic (Def. 5.3) and that moreover the map (5.5) can be chosen such that*

$$f_t(x)\neq y \quad \text{for} \quad x\subset B, \quad 0\leq t\leq 1, \tag{5.7}$$

Then

$$j(f,y,D)=j(g,y,D). \tag{5.8}$$

PROOF. It will be obviously sufficient to show that $j(f_t,y,D)$ is continuous in t. To do this we first prove

LEMMA 5.1[30]. *For each t in the unit interval let $f_t(x)=f(x,t)=x-F(x,t)$ be a map of $\overline{D}$ into X which satisfies conditions* i) *and* ii) *of definition* 5.1, *is continuous on the product of $\overline{D}$ with the unit interval, and satisfies conditions* a) *and* b) *of definition* 5.3. *Let ϑ, ϑ_1, a_0, b_0, and τ be positive numbers which are chosen as follows:*

$$0<\vartheta<\vartheta_1<1. \tag{5.9}$$

a_0 is then chosen in such a way that

$$\|d(x,t,h)-d(x_0,t,h\|<\|h\|\vartheta/m \quad \text{for} \quad \|x-x_0\|<a_0 \tag{5.10}$$

where m is a number satisfying (5.6); this choice is possible by assumption i) *of definition* 5.1. *For b_0 we then choose any number satisfying*

$$0<b_0<a_0(1-\vartheta_1)/m. \tag{5.11}$$

Let x_0 be a point of D and set

$$f(x_0,t)=y_0(t) \qquad (0\leq t\leq 1). \tag{5.12}$$

[30]) Lemma 5.1 and its proof are essentially an adaptation to the present purpose of results and methods contained in [3]. The proof is here given for the sake of completeness and of the possibility of referring later on to some of the inequalities contained in it.

Then choose $\tau > 0$ in such a way that

$$\|y_0(t) - y_0(t_0)\| < b_0/2 \quad \text{for} \quad |t - t_0| < \tau. \tag{5.13}$$

Finally let y be any point of X which satisfies

$$\|y - y_0(t)\| < b_0 \quad \text{for} \quad |t - t_0| < \tau. \tag{5.14}$$

Then the equation for x

$$y = f(x, t) \tag{5.15}$$

has one and only one solution $x = x(y, t)$ which satisfies

$$\|x(t, y) - x_0\| < a_0 \qquad (|t - t_0| < \tau) \tag{5.16}$$

and this solution is continuous in the product $D \times \{|t - t_0| < \tau\}$.

PROOF. Subtracting (5.12) from (5.15) and setting

$$\begin{cases} x - x_0 = \xi, \quad y - y_0(t) = \eta, \\ f(x, t) - f(x_0, t) - d(x_0, t, x - x_0) = R(x, x_0, t) \end{cases} \tag{5.17}$$

we see that the equation for ξ

$$\eta(t) = d(x_0, t, \xi) + R(x_0 + \xi, x_0, t) \tag{5.18}$$

is equivalent to (5.15). Multiplying (5.18) by $d^{-1} = l$ and setting

$$l(\eta) = \zeta, \quad l\big(R(x_0 + \xi, x_0, t)\big) = -S(x_0, t, \xi) \tag{5.19}$$

we obtain for ξ the equation

$$\xi = \zeta(t) + S(x_0, t, \xi) \tag{5.20}$$

which is still equivalent to (5.15). In order to solve (5.20) by successive approximations we prove the following two inequalities:

$$\|\zeta\| < a_0(1 - \vartheta_1) \tag{5.21}$$

if (5.14) is satisfied, and

$$\|S(x_0, t, \xi'') - S(x_0, t, \xi')\| < \vartheta \|\xi'' - \xi'\| \tag{5.22}$$

if

$$\|\xi''\| < a_0, \quad \|\xi'\| < a_0. \tag{5.23}$$

The proof of (5.21) is immediate since by (5.19), (5.6), (5.17), (5.14), and (5.11)

$$\|\zeta\| < m\|\eta\| = m\|y - y_0(t)\| < m\, b_0 < a_0(1 - \vartheta_1).$$

As to the proof of (5.22) we obtain from the definition (5.17) of R using the fact that $d(x, t, \xi)$ is linear in its third argument

$$\begin{cases} R(x_0 + \xi'', x_0, t) - R(x_0 + \xi', x_0, t) \\ \qquad = [f(x_0 + \xi'', t) - f(x_0 + \xi', t)] - d(x_0, t, \xi'' - \xi'). \end{cases} \tag{5.24}$$

If we apply the mean value theorem[31]) to the difference in the bracket, the right member of (5.24) may be written as

$$\int_0^1 \{d(x_0+\xi'+\beta(\xi''-\xi'), t, \xi''-\xi') - d(x_0, t, \xi''-\xi')\}\, d\beta.$$

Since (5.23) implies the corresponding inequality for $\xi'+\beta(\xi''-\xi')$ if $0\leq\beta\leq 1$, we see from (5.10) (with h replaced by $\xi''-\xi'$) that the norm of the above integral is $<\|\xi''-\xi'\|\vartheta/m$. Therefore using (5.6) and the definition (5.19) of S we see that

$$\begin{aligned}\|S(x_0, t, \xi'') - S(x_0, t, \xi')\| &= \|l\{R(x_0+\xi'', x_0, t) - R(x_0+\xi', x_0, t)\}\| \\ &< m\|R(x_0+\xi'', x_0, t) - R(x_0+\xi', x_0, t)\| \\ &< m\|\xi''-\xi'\|\vartheta/m\end{aligned}$$

which proves (5.22). As a special case we note that the inequality

$$\|S(x_0, t, \xi)\| < \vartheta\|\xi\| \tag{5.25}$$

is obtained from (5.22) by setting $\xi''=\xi$, $\xi'=0$ since $S(x_0, t, 0)=0$ as is seen from (5.19) and (5.17).

Now from the inequalities (5.21), (5.22), and (5.25) one proves in the well-known fashion (see, e.g., [3]) that the iterations

$$\xi_1=\zeta, \quad \xi_{n+1}=\zeta+S(x_0, t, \xi_n) \qquad (n=1, 2, \ldots)$$

are well defined (for ζ and t in the domains specified) and satisfy the inequalities

$$\|\xi_n\| < a_0(1-\vartheta_1)/(1-\vartheta), \quad \|\xi_n-\xi_{n-1}\| < a_0\vartheta^{n-1}$$

which show that $\xi=\xi(\zeta, t)=\lim_{n\to\infty}\xi_n$ exists uniformly and that

$$\|\xi\| \leq a_0(1-\vartheta_1)/(1-\vartheta) < a_0.$$

Obviously, ξ satisfies (5.20), and $x=x(y, t)$ obtained from $\xi=\xi(\zeta, t)$ by (5.17) and (5.19) satisfies (5.15). The continuity of $x(y, t)$ follows from the uniformity of the convergence since each iteration ξ_n is continuous in (ζ, t). The uniqueness of ξ as solution of (5.20) follows in the usual way from (5.22) and (5.9).

We now return to the proof of theorem 5.1. We have to show: if $y=y_1$ is a point of X for which (5.7) is satisfied then there exists a $\tau>0$ such that

$$j(f_t, y_1, D) = j(f_{t_0}, y_1, D) \quad \text{for} \quad |t-t_0|<\tau. \tag{5.26}$$

We distinguish two cases:

Case 1. f_{t_0} maps no point of D on y_1. Then by (5.7) no point of the closed set $\overline{D}$ is mapped on y_1 by f_{t_0}, and it follows from assumption v) of definition 5.1 that the image under f_{t_0} of $\overline{D}$ has a positive distance from y_1. This again implies by definition 5.3 that for small enough $|t-t_0|$, y_1 is not image point of $\overline{D}$ under the map f_t. Therefore for such t (5.26) is true, both members of the equation being zero by definition 5.2.

[31]) [2].

CASE 2. y_1 is image point of D under the map $f_{t_0}=f(x,t_0)$. Then by assumption iii) of definition 5.1 there are only a finite number of points in D, say $x^1, x^2, \ldots, x^r$ in D such that

$$f(x^\varrho, t_0) = y_1 \qquad (\varrho = 1, 2, \ldots, r). \tag{5.27}$$

Now by theorem 2.1 there exists an $\varepsilon > 0$ such that any bounded linear operator $l^\varrho(\xi)$ which satisfies the inequality

$$\|d(x^\varrho, t_0, \xi) - l^\varrho(\xi)\| < \varepsilon\|\xi\| \tag{5.28}$$

satisfies also the equality

$$j(d^\varrho, t_0, \xi) = j\left(l^\varrho(\xi)\right). \tag{5.29}$$

In order to construct the required τ we remark first that by assumption a) of definition 5.3 it is possible to choose two positive numbers τ_1 and δ such that

$$\|d(x^\varrho, t_0, \xi) - d(x, t, \xi)\| < \varepsilon\|\xi\| \tag{5.30}$$

if

$$|t - t_0| < \tau_1 \tag{5.31}$$

and

$$\|x - x^\varrho\| < \delta. \tag{5.32}$$

We now choose a number a_0 which satisfies the following requirements: i) (5.10) is satisfied with $x_0 = x^\varrho$; ii) we have

$$|a_0| < \delta; \tag{5.33}$$

iii) the r solid spheres with centers x^ϱ and radius a_0 are disjoint and contained in D. We then choose a number b_0 subject to the restriction (5.11), and a number τ_2 such that

$$\|f(x^\varrho, t) - f(x^\varrho, t_0\| < b_0/2 \tag{5.34}$$

if

$$|t - t_0| < \tau_2. \tag{5.35}$$

We finally choose a positive number τ_3 as follows: let C be the closure of the set obtained from D by cutting out the spheres with centers x^ϱ and radius a_0, and denote by $f_t(C)$ the image of C under the map f_t. Then $f_{t_0}(C)$ does not contain y_1. Therefore by assumption v) of definition 5.1, $f_{t_0}(C)$ has a positive distance from y_1. Consequently, there exists a τ_3 such that $f_t(C)$ does not contain y_1 for $|t-t_0|<\tau_3$; we can then conclude that for some ϱ of the set $1, 2, \ldots, r$

$$\|x - x^\varrho\| < a_0 \quad \text{if} \quad f(x,t) = y_1 \quad \text{and} \quad |t - t_0| < \tau_3. \tag{5.36}$$

We claim that (5.26) holds with $\tau = \min(\tau_1, \tau_2, \tau_3)$. Indeed, consider $f_t = f(x,t)$ for a t in $|t-t_0|<\tau$. Let

$$f(x^\varrho, t) = y^\varrho(t). \tag{5.37}$$

With the choic of the constants a_0, b_0, τ made we can apply lemma 5.1 with $x_0 = x^\varrho$ and $y_0(t) = y^\varrho(t)$ and see that the equation

$$f(x, t) = y \tag{5.38}$$

has for all y satisfying

$$\|y - y^\varrho(t)\| < b_0 \tag{5.39}$$

one and only one solution $x^\varrho(y, t)$ which satisfies the inequality

$$\|x^\varrho(y, t) - x^\varrho\| < a_0. \tag{5.40}$$

Now by (5.34), (5.27), and (5.37), $y = y_1 = f(x^\varrho, t_0)$ satisfies (5.39). Therefore the equation for x

$$f(x, t) = y_1 \tag{5.41}$$

has one and only one solution $x_1^\varrho = x^\varrho(y_1, t)$ in (5.40). Then, by (5.33), $x = x_1^\varrho$ satisfies (5.32). Consequently, since t satisfies (5.31) we see that $d(x_1^\varrho, t, \xi)$ satisfies (5.30), i.e., that (5.28) is true with $l^\varrho(\xi) = d(x_1^\varrho, t, \xi)$, and gives us the equality

$$j\left(d(x^\varrho, t_0, \xi)\right) = j\left(d(x_1^\varrho, t, \xi)\right). \tag{5.42}$$

Since by (5.36) $x_1^1, x_1^2, \ldots x_1^r$, are the only solutions in D of (5.41) we obtain by definition 5.2 of the degree the desired equality (5.26) by summing (5.42) over ϱ.

THEOREM 5.2. *Let $y = f(x)$ be a smooth mapping of the closure of the open bounded domain D of X into X, and y_0 a point of X which is not situated on the image of the boundary B of D. Then there exists a positive number b_0 such that*

$$j(f, y, D) = j(f, y_0, D) \quad \text{for} \quad \|y - y_0\| < b_0. \tag{5.43}$$

PROOF. We detinguish two cases. Case 1. No point of D is mapped on y_0. Then it is easily seen from assumption v) of definition 5.1 that for small enough $\|y - y_0\|$, the point y is also not in the image of D or B. Therefore for such y equation (5.43) is true, both members of the equality being zero.

CASE 2. Some points of D are mapped on y_0. In this case let $x_0^1, x_0^2, \ldots, x_0^r$ be the points of D which are mapped on y_0. We choose a number a_0 which satisfies the following conditions: i) (5.10) is satisfied with $x_0 = x_0^\varrho$ $(\varrho = 1, 2, \ldots, r)$; ii) the equality

$$j\left(d(x_0^\varrho, \xi)\right) = j\left(d(x, \xi)\right) \quad \text{for} \quad \|x - x_0^\varrho\| < a_0 \tag{5.44}$$

holds for $\varrho = 1, 2, \ldots, r$, this choice being possible on account of (5.2) and theorem 2.1; iii) the spheres with centers x_0^ϱ and radius a_0 are disjoint and lie in D. We then choose a number b_0 subject to the following conditions: i) (5.11) is satisfied; ii) the sphere with radius b_0 about y_0 does not intersect the image of B which choice is possible by assumption v) of definition 5.1; iii) let C be the closure of the complement in D of the union of the spheres $\|x - x_0^\varrho\| < a_0$; the image of C does not contain y_0 and has therefore (by the

assumption just mentioned) a positive distance from y_0; we choose b_0 less than this distance. We claim that (5.43) is satisfied with this choice of b_0. Indeed, if $\|y-y_0\|<b_0$ then the equation $f(x)=y$ can have a solution only in one of the spheres $\|x-x_0^\varrho\|<a_0$, and by lemma 5.1[32]) it has exactly one, say $x=x^\varrho$, in each of these spheres. Then (5.44) holds for $x=x^\varrho$, and it follows from definition 5.2 that we obtain (5.43) by summing (5.44) (with x replaced by x^ϱ) over ϱ.

6. REAL BANACH SPACES. The preceding sections dealt with complex Banach spaces. However, it is easy to carry over the theory from such spaces to real Banach spaces by imbedding them into a complex Banach space as will be done in this section.

DEFINITION 6.1. We say that the real Banach space R is imbedded in the complex Banach space X if X has a conjugation (def. 3.1) of the following property: let R_1 be the Banach space over the reals which is formed by the elements of X which are invariant under this conjugation[33]); then R and R_1 are algebraically isomorphic and isometric.

LEMMA 6.1. *Every real Banach space R can be imbedded in a complex Banach space.*

PROOF. We construct X as follows: the elements x of X are the ordered couples $x=(x', x'')$ where x', x'' are elements of R. If $x=(x', x'')$ and $y=(y', y'')$ are two elements of X and $\alpha=\alpha'+i\,\alpha''$ a complex number we define the linear operations by

$$x+y=(x'+y', x''+y''), \qquad \alpha\,x=(\alpha'\,x'-\alpha''\,x'', \alpha'\,x''+\alpha''\,x').$$

Finally, following Krein and Gelfand [I. Gelfand, Normierte Ringe, Recueil Math., T. 9 (51), 1941, 3—23, § 8] we define the norm $\|x\|$ of the element $x=(x', x'')$ by $\|x\|=\sup_\varphi\|x'\cos\varphi+x''\sin\varphi\|$.

DEFINITION 6.2. Let

$$y=f(x)=x-F(x) \tag{6.1}$$

be a linear non singular map of the real Banach space R into itself with completely continuous $F(x)$. Denote by $\lambda_1, \lambda_2, \ldots, \lambda_p$ those eigenvalues of $F(x)$ which are real greater than 1[34]). Let m_i be the number of linearly independent solutions in R of the totality of equations $(\lambda_i x-F(x))^n=0$ for $n=1, 2, \ldots$[35]). We then define the index $j(f)$ of f by the equation

$$j(f)=\sum_{i=1}^{p}(-1)^{m_i}. \tag{6.2}$$

---

[32]) We apply lemma 5.1 with $f(x, t)\equiv f(x)$.

[33]) Cf. Definition 3.2 and lemma 3.1.

[34]) That there are at most a finite number of eigenvalues >1 is a well known consequence of the complete continuity of F.

[35]) It is well known that m_i is finite.

THEOREM 6.1. *Let $F(x)$ satisfy the assumptions of definition* 6.2. *Then there exists a $\delta>0$ of the following property: if $g(x)=x-G(x)$ where $G(x)$ is completely continuous and*

$$\|f-g\|=\|F-G\|<\delta, \tag{6.3}$$

then

$$j(f)=j(g). \tag{6.4}$$

PROOF. Let X be the complex Banach space in which R is imbedded by lemma 6.1. By lemma 3.2 every element of x of X is of the form

$$x=x'+ix'' \quad \text{where} \quad x'=(x+\bar{x})/2 \quad \text{and} \quad x''=(x-\bar{x})/2i \text{ are in } R. \tag{6.5}$$

We define

$$\tilde{F}(x)=F(x')+iF(x''), \quad \tilde{f}(x)=x-\tilde{F}(x)$$

and $\tilde{G}$, $\tilde{g}$ correspondingly. It is then easily verified that $\tilde{F}(x)$ is a linear completely continuous map of X into itself and that the real points of its spectrum coincide with the real eigenvalues of $F(x)$. Moreover, $\tilde{F}(x)$ is real in the sense of definition 3.3 and satisfies the assumptions A', B', C' of section 3 (cf. remark 3.1). Now by remark 4.2 the numbers m_i of definition 6.1 agree with those of theorem 3.1. Therefore

$$j(f)=j(\tilde{f}) \tag{6.6}$$

and similarly

$$j(g)=j(\tilde{g}). \tag{6.7}$$

Now by theorem 3.2 there exists an $\varepsilon>0$ such that

$$j(\tilde{f})=j(\tilde{g}) \tag{6.8}$$

if

$$\|\tilde{F}-\tilde{G}\|<\varepsilon. \tag{6.9}$$

We now set $\delta=\varepsilon/2$ and claim that with this δ, (6.3) implies (6.4). Indeed, if we set $H=F-G$ and $\tilde{H}=\tilde{F}-\tilde{G}$, we obtain, recalling (6.5) and (3.3)

$$\begin{aligned}\|\tilde{F}(x)-\tilde{G}(x)\|&=\|\tilde{H}(x)\|=\|H(x')+iH(x'')\|\leq\|H\|\{\|x'\|+\|x''\|\}\\&=\|H\|\{\|x+\bar{x}\|+\|x-\bar{x}\|\}/2\\&\leq\|H\|\{\|x\|+\|\bar{x}\|\}=2\|H\|\cdot\|x\|\\&=2\|F-G\|\|x\|\leq 2\delta\|x\|=\varepsilon\|x\|\end{aligned}$$

thus proving (6.9) and therefore (6.8). (6.4) follows now from (6.6) and (6.7).

REMARK 6.1. Let R be a real Banach space. Let D be a bounded open set of R, $\bar{D}$ its closure, B its boundary, and let (6.1) be a map of $\bar{D}$ into R with completely continuous (not necessarily linear) $F(x)$. We then call $f(x)$ smooth if it satisfies conditions i) and ii) of definition 5.1 and if in addition the linear operator in h, $D(x,h)=h-d(x,h)$ is completely continuous (cf.

Remark 5.1). Perusal of section 5.1 shows then that everything said in this section remains valid if the following substitutions are made throughout: i) X is replaced by R; ii) definition 5.1 of a smooth map is replaced by the definition just given; iii) references to theorem 2.1 are replaced by references to theorem 6.1, and definition 2.1 by definition 6.2.

By the preceding remark the degree is defined and homotopy properties are established for smooth mappings in a real Banach space R[36]).

Bibliography.

[1] DUNFORD, N.: Spectral theory. I. Convergence to projections. Trans. Amer. Math. Soc. **54**, 185—217 (1943). — [2] GRAVES, L. M.: Riemann integration and Taylor's theorem in general analysis. Trans. Amer. Math. Soc. **29**, 163—177 (1927). — [3] GRAVES, L. M., and T. H. HILDEBRANDT: Implicit functions and their differentials in general analysis. Trans. Amer. Math. Soc. **29**, 126—153 (1927). — [4] HILLE, E.: Functional analysis and semi-groups. Amer. Math. Soc., Colloquium Publications **21** (1948). — [5] LERAY, J.: La theorie des points fixes et ses applications en analyse. Proc. Internat. Congress of Mathematicians, Cambridge, Mass. U.S.A., 1950. Publ. Amer. Math. Soc. **2**, 202—207 (1952). — [6] LERAY, J., et J. SCHAUDER: Topology and équations fonctionnelles. Ann. sci. École norm. sup. (3) **51**, 45—78 (1934). — [7] MACKEY, G. W.: Isomorphisms of normed linear spaces. Ann. Math. **43**, 244—260 (1942). — [8] NAGUMO, M.: A theory of the degree of mapping based on infinitesimal analysis. Amer. J. Math. **73**, 485—496 (1951). — [9] NAGUMO, M.: Degree of mapping in convex linear topological spaces. Amer. J. Math. **73**, 485—496 (1951). — [10] SCHWARTZ, J.: Perturbations of spectral operators and applications. I. Bounded perturbations. Pacific J. Math. **4**, 415—458 (1954). — [11] STONE, M. H.: Linear transformations in Hilbert space and their applications to analysis. Amer. Math. Soc., Colloquium Publications **15** (1932).

[36]) This degree agrees with the one defined by Leray and Schauder. The latter, however, applies to a wider class of mappings.

Dept. of Math., University of Michigan, Ann Arbor, Mich. (U.S.A.).

(Eingegangen am 16. März 1955.)

BERNAYS, P.
Math. Zeitschr. Bd. 63, S. 219—229 (1955)

Betrachtungen über das Vollständigkeitsaxiom und verwandte Axiome.

Von

PAUL BERNAYS.

Meinem verehrten Lehrer ISSAI SCHUR zum Gedenken.

§ 1. Die im folgenden dargestellten Überlegungen knüpfen sich an die Formulierung der Stetigkeitseigenschaft des Systems der reellen Zahlen mit Hilfe des Hilbertschen Vollständigkeitsaxioms. Dieses Axiom, an den Schluß eines Axiomensystems gestellt, besagt ja, als „Maximalaxiom", daß das zu kennzeichnende System von Dingen bei Aufrechterhaltung der Beziehungen und Verknüpfungen zwischen diesen Dingen sowie auch der Erfüllung der vorhergehenden Axiome keiner Erweiterung fähig ist.

Der methodisch komplizierte Charakter dieser Forderung, wie ihn die Bezugnahme auf die anderen Axiome mit sich bringt[1]), läßt sich jeweils beheben; jedoch muß dann im allgemeinen eine schwerfällige Fassung und eine Einbuße an Prägnanz in Kauf genommen werden. Es soll nun zunächst gezeigt werden, daß im Falle des Axiomensystems für die reellen Zahlen das Vollständigkeitsaxiom von der Bezugnahme auf die anderen Axiome auf eine ungezwungene Art und ohne Verlust an Prägnanz befreit werden kann.

Hierzu gehen wir davon aus, daß bei dem betrachteten Axiomensystem[2]) die dem Vollständigkeitsaxiom vorhergehenden Axiome besagen, daß die reellen Zahlen mit ihren Rechenverknüpfungen und der Größenbeziehung einen geordneten Körper bilden, welcher dem Archimedischen Axiom genügt, kurz: einen archimedischen geordneten Körper. Von der Ordnung eines Körpers soll eo ipso verlangt sein, daß außer den allgemeinen Eigenschaften der Beziehung $<$ auch die beiden folgenden zutreffen:

(1) Wenn $a<b$ und c beliebig, so $a+c<b+c$.

(2) Wenn $0<a$ und $0<b$, so $0<a\cdot b$.

Eine Erweiterung eines Körpers mit Erhaltung der Rechenverknüpfungen und ihrer Gesetze wird üblichermaßen eine *Körpererweiterung* genannt. Die Aussage des Vollständigkeitsaxioms bezogen auf die Axiome, welche einen

[1]) Auf diese speziell für die axiomatischen Untersuchungen sich auswirkende Kompliziertheit wurde insbesondere von R. BALDUS hingewiesen: Zur Axiomatik der Geometrie. I. Math. Ann. **100**, 321—333 (1928). Daß andererseits diese Komplikation die logistische Formalisierung des Vollständigkeitsaxioms nicht hindert, wurde von R. CARNAP und F. BACHMANN gezeigt: Über Extremalaxiome. Erkenntnis **6**, 166—188 (1936).

[2]) Vgl. die Fassung im Anhang VI („Über den Zahlbegriff") der „Grundlagen der Geometrie" von D. HILBERT, 7. Aufl. 1930.

archimedischen geordneten Körper als solchen kennzeichnen, ist nun ersichtlich gleichbedeutend mit der Forderung, daß jede Körpererweiterung, welche zugleich eine Erweiterung der Ordnung ist, zu einem nicht-archimedischen Körper führt. Demnach erhalten wir die folgende Charakterisierung des Systems der reellen Zahlen: *Das System der reellen Zahlen ist ein archimedischer geordneter Körper von der Eigenschaft, daß jeder geordnete Erweiterungskörper, der die vorherige Ordnung beibehält, nicht-archimedisch ist.*

Wir können hier die Formulierung noch enger wählen, ohne an dem axiomatisch Geforderten etwas zu ändern. Nämlich es genügt ja, jeweils solche Erweiterungskörper zu betrachten, die (nach der üblichen Ausdrucksweise der Algebra) durch Adjunktion eines einzigen Elementes zu dem betreffenden Körper entstehen. Denn jeder Erweiterungskörper hat ja einen Körper der genannten Art als Unterkörper. Somit läßt sich die obige Formulierung durch die folgende ersetzen: *Das System der reellen Zahlen ist ein solcher archimedischer geordneter Körper, bei welchem jede Erweiterung durch Adjunktion eines neuen Elementes entweder die Ordnung stört oder zu einem nicht-archimedischen Körper führt.*

Beachten wir ferner, daß jede *algebraische* Erweiterung eines archimedischen geordneten Körpers, welche die Ordnung erhält, wieder einen archimedischen geordneten Körper ergibt, so gelangen wir zu folgender Formulierung: *Das System der reellen Zahlen ist ein solcher archimedischer geordneter Körper, bei welchem 1. jede algebraische Erweiterung die Ordnung stört, 2. jede transzendente Erweiterung nur mit einer solchen Erweiterung der Ordnung verbunden werden kann, welche nicht-archimedisch ist.*

Die Bedingung 1. ist gleichwertig derjenigen, daß der Körper, wie man es nennt, reell algebraisch abgeschlossen ist; dieses kann man z.B. daraus ersehen, daß einerseits in einem Körper, der diese Bedingung erfüllt, jedes positive Element eine Quadratwurzel besitzen muß und daher, für jedes Element a des Körpers, a selbst oder $-a$ ein Quadrat ist, und daß andererseits bei einem Körper mit dieser Eigenschaft eine den Anforderungen (1) und (2) entsprechende Ordnung nur auf eine Art möglich ist. Die Bedingung 2. besagt, daß bei einer Erweiterung durch Adjunktion eines relativ transzendenten Elementes dieses neue Element sich nicht archimedisch unter die vorherigen Elemente einordnen läßt.

§ 2. Die Situation, wie sie bei einem archimedischen geordneten und reell abgeschlossenen Körper K vorliegt, wenn jede Adjunktion eines transzendenten Elementes a den archimedischen Charakter der Ordnung aufhebt, läßt sich näher verdeutlichen.

Der aus K durch Adjunktion von a entstehende Körper $K(a)$ sei in solcher Weise geordnet, daß die Bedingungen (1) und (2) erfüllt sind und daß für die Elemente von K die Ordnung die vorherige bleibt. Schalten wir nun den Fall aus, wo a „unendlich groß" ist, d.h. entweder größer als jedes ganzzahlige positive oder kleiner als jedes ganzzahlige negative Vielfache der Einheit e von K, so erzeugt a einen Schnitt in der Menge der rationalen Vielfachen

von e. Würde dieser Schnitt nicht auch durch ein Element von K erzeugt, so gäbe es in einem archimedischen Erweiterungskörper K' ein Element, das ihn erzeugt. Dieses Element könnte aber nach unseren Annahmen über K weder algebraisch in bezug auf K noch auch relativ transzendent zu K sein. Somit gibt es in K ein Element b, das den gleichen Schnitt in der Menge der rationalen Vielfachen von e erzeugt, wie das zu K adjungierte Element a; und es kann daher zwischen a und b kein Element des archimedischen Körpers K liegen. Die aus den Elementen von K und dem Element a gebildete Menge in der im Körper $K(a)$ gegebenen Ordnung ist somit *nicht dicht geordnet*.

So ergibt sich, daß bei einem archimedischen geordneten Körper K, welcher die am Ende von § 1 formulierte Vollständigkeitsbedingung erfüllt, jede Erweiterung zu einem geordneten Körper $K(a)$ so beschaffen ist, daß die geordnete, aus a und den Elementen aus K bestehende Teilmenge von $K(a)$ entweder a als erstes oder letztes Element hat, oder zwei unmittelbar aufeinanderfolgende Elemente besitzt, und daher nicht dicht geordnet ist [3]). Diese Feststellung legt es nahe, in der axiomatischen Charakterisierung des Systems der reellen Zahlen die körpertheoretische Vollständigkeitsbedingung durch die einfachere ordnungstheoretische Bedingung zu ersetzen, daß bei Hinzunahme eines Elementes keine solche Erweiterung der Ordnung möglich ist, bei welcher die Eigenschaft der Dichtigkeit der Ordnung zusammen mit derjenigen, daß kein erstes und kein letztes Element vorhanden ist, aufrechterhalten wird.

Für einen geordneten Körper ist diese Bedingung, wie man ohne Mühe erkennt, gleichbedeutend mit derjenigen, daß die Ordnung eine *im Dedekindschen Sinne lückenlose* ist, worin zugleich enthalten ist, daß sie eine archimedische ist. Andererseits existiert, wie man weiß, zu jedem archimedischen, aber nicht lückenlos geordneten Körper ein archimedischer Erweiterungskörper.

Auf diese Weise ergibt sich, daß einerseits bei Verwendung der ordnungstheoretischen Vollständigkeitsbedingung die Forderung, daß der Körper archimedisch sei, weggelassen werden kann, und andererseits, daß in Anwendung auf archimedische Körper die ordnungstheoretische Vollständigkeitsbedingung nicht mehr verlangt als die (dem Wortlaut nach schwächere) körpertheoretische Vollständigkeitsbedingung. Demnach läßt sich *das System der reellen Zahlen axiomatisch auch charakterisieren als ein solcher geordneter Körper, dessen Ordnung bei Hinzunahme eines Elementes entweder die Eigenschaft der Dichtigkeit einbüßt oder ein erstes oder ein letztes Element erhält.*

Daß in dieser Kennzeichnung die Forderung des archimedischen Charakters der Ordnung als herleitbar wegfällt, bedeutet einerseits eine Vereinfachung, kann jedoch andererseits in Hinsicht auf die axiomatische Analyse als Nachteil empfunden werden.

§ 3. Zwischen der körpertheoretischen und der ordnungstheoretischen Fassung der Vollständigkeitseigenschaft des Systems der reellen Zahlen kann

[3]) Man beachte, daß gleichwohl der Körper $K(a)$, sowie jeder geordnete Körper, dicht geordnet ist.

ein Mittelweg gesucht werden, auf eine Art, die keineswegs künstlich ist. Dazu werden wir hingeleitet, wenn wir erwägen, worauf allgemein das Prädikat „archimedisch" zu beziehen ist. Daß dieses Prädikat nicht rein ordnungstheoretisch ist, wird schon dadurch plausibel gemacht, daß in der Formulierung des Archimedischen Postulates der nicht-ordnungstheoretische Begriff des Vielfachen auftritt. Dieser Umstand allein würde aber noch nicht ausschließen, daß etwa jeder archimedische Körper schon dem Ordnungstypus nach von einem nicht-archimedischen verschieden wäre. Daß aber dieses tatsächlich nicht der Fall ist, zeigt das einfachste Beispiel eines nicht-archimedischen Körpers: der Körper aller derjenigen rationalen Funktionen von t, welche ausgehend von der Variablen t durch Anwendung der elementaren Rechenoperationen erhalten wird, wobei das Vorzeichen und damit die Ordnung sich folgendermaßen bestimmt. Jedes Element außer 0 ist darstellbar in der Form

$$t^k \cdot \frac{a_0 + a_1 t + \cdots + a_m t^m}{1 + b_1 t + \cdots + b_n t^n},$$

wobei k ganz rational, $a_0, a_1 \ldots a_m, b_1, b_2 \ldots b_n$ rational und $a_0 \neq 0$ ist. Als Vorzeichen des Elementes wird das Vorzeichen von a_0 genommen und die rationale Funktion $P(t)$ heißt größer als $Q(t)$, kleiner als $Q(t)$ oder gleich $Q(t)$, je nachdem $P(t) - Q(t)$ positiv, negativ oder 0 ist. Diese Ordnung ist eine dichte und ohne erstes und letztes Element. Außerdem aber lassen sich die Elemente des Körpers in eine Abzählung bringen. Somit ist der Typus der Ordnung nach einem Satz der Mengenlehre der gleiche wie derjenige der gewöhnlichen Ordnung des archimedischen Körpers der rationalen Zahlen.

Der Unterschied zwischen dem Archimedischen und Nicht-archimedischen bezieht sich also nicht schlechtweg auf die Ordnung. Andererseits ist die Unterscheidung nicht nur auf geordnete Körper anwendbar, sondern auch schon auf geordnete additive Gruppen, bei denen die Ordnung mit der Addition durch die Bedingung (1) verkoppelt ist. Eine solche geordnete Gruppe ist vollständig bestimmt durch das System ihrer positiven Elemente. Diese bilden ein *Größensystem*. Darunter verstehen wir eine additive Halbgruppe ohne 0-Element und mit relativen Komplementen, genauer gesagt: ein System von Elementen, für welche eine assoziative und kommutative Addition definiert ist mit den Eigenschaften, daß $a + b \neq a$ und für $a \neq b$ stets eine der Gleichungen $a + x = b$, $b + x = a$ eine Lösung x hat. Hier läßt sich die Ordnungsbeziehung $a < b$ durch die Lösbarkeit der Gleichung $a + x = b$ definieren, und es werden dann alle Eigenschaften der Größenbeziehung beweisbar.

Von einem Größensystem kann man stets durch Bildung von Paaren und geeignete Definitionen von Gleichheit, Kleinerbeziehung und Summe zu einer geordneten additiven Gruppe übergehen, für welche das System der positiven Elemente jenem Größensystem isomorph ist. Das geschieht ganz analog wie der Übergang von einem Ring ohne Nullteiler zu dem Quotientenkörper.

Die Größensysteme sind nun diejenigen geordneten Gesamtheiten, auf welche sich ursprünglich der Unterschied von Archimedisch und Nicht-archimedisch bezieht. Hier können wir auch einen Vollständigkeitsbegriff

im Sinne des Vollständigkeitsaxioms einführen. Ein archimedisches Größensystem möge vollständig heißen, wenn jedes echt umfassende Größensystem nicht-archimedisch ist. Man zeigt auf bekannte Weise, daß ein Größensystem dann und nur dann ein archimedisches und vollständiges ist, wenn seine Ordnung eine dichte und lückenlose ist.

Für die archimedischen Größensysteme läßt sich auf klassische Art die Proportionenlehre mittels Dedekindscher Schnitte gewinnen, indem das Verhältnis $c:a$ erklärt wird als die Menge derjenigen geordneten Paare natürlicher Zahlen $\langle k, l\rangle$, für welche der Wert der k-gliedrigen Summe aus Summanden a kleiner ist als der Wert der l-gliedrigen Summe aus Summanden c. Die Proportion $c:a=d:b$ wird dann durch die Übereinstimmung der Mengen $c:a$, $d:b$ erklärt und die Ungleichung $c:a<d:b$ als die Aussage, daß die Menge $c:a$ eine echte Teilmenge von $d:b$ ist. Die Paarmengen, welche die Verhältnisse bilden, besitzen die bekannten Eigenschaften eines Schnittes. Im Falle der Vollständigkeit des Größensystems läßt sich andererseits beweisen, daß jede Menge von Paaren $\langle k, l\rangle$, welche diese Eigenschaften besitzt, als Verhältnis von Größen auftritt, sowie auch, daß zu drei Größen stets eine vierte Proportionale existiert.

Von diesen Begriffen ausgehend, kann nun das *System der positiven reellen Zahlen gekennzeichnet werden als ein vollständiges archimedisches Größensystem, worin ein Element e als Einheit ausgezeichnet ist.* Die Multiplikation läßt sich hier definitorisch einführen, indem das Produkt von a und b erklärt wird als dasjenige Element c, welches die Proportion $c:a=b:e$ erfüllt. Von einem solchen Größensystem kann man auf die vorhin erwähnte Art zu einer additiven Gruppe übergehen, womit dann die Struktur des Systems aller reellen Zahlen gewonnen wird. *Das System der reellen Zahlen kann also auch gekennzeichnet werden als diese Erweiterung eines vollständigen archimedischen Größensystems mit ausgezeichneter Einheit zu einer additiven Gruppe.*

Es sei noch darauf hingewiesen, daß die Beweisbarkeit des archimedischen Charakters einer Ordnung aus der Lückenlosigkeit nur für die Ordnung von Größensystemen oder von additiven Gruppen besteht. Allgemein kann ja ein jeder Ordnungstypus durch geeignete Hinzufügung von Elementen zu einem lückenlosen ergänzt werden, jedoch gelangt man durch die Anwendung dieses Verfahrens auf die Ordnung eines nicht-archimedischen Größensystems nicht wieder zu einem Größensystem.

§ 4. Wie schon am Ende von § 2 erwähnt wurde, erscheint es vom axiomatischen Standpunkt als wünschenswert, die Formulierung derjenigen Stetigkeitseigenschaft eines Größensystems oder eines Körpers, welche zu der im Archimedischen Axiom ausgesprochenen hinzutritt, von diesem Axiom abzusondern. Dieser Zweck wird bei der Anwendung des Axioms der Lückenlosigkeit nicht erreicht, da ja aus der Lückenlosigkeit eines Größensystems schon folgt, daß dieses archimedisch ist. Aber auch die in § 1 und § 3 formulierten Vollständigkeitsbedingungen sind in der genannten Hinsicht nicht befriedigend, da diese Forderungen für nicht-archimedische Größensysteme

bzw. Körper (abgesehen von derjenigen der reell algebraischen Abgeschlossenheit), eo ipso erfüllt und insofern doch nur als Ergänzungen zum Archimedischen Axiom bedeutungsvoll sind.

Eine wirkliche Absonderung der zusätzlichen Vollständigkeitsbedingung von dem Archimedischen Axiom läßt sich aber erreichen. Insbesondere gelingt dieses mittels des *Cantorschen Stetigkeitsaxioms*, welches in seiner rein ordnungstheoretischen Form besagt: Zu jeder Folge von Intervallen (Elementepaaren) (a_n, b_n) von der Eigenschaft, daß

$$a_n < a_{n+1} < b_{n+1} < b_n \quad (n = 1, 2, \ldots),$$

gibt es ein Element c, das in sämtlichen Intervallen liegt. Die Forderung dieses Axioms ist aus der Forderung der Lückenlosigkeit als Konsequenz zu entnehmen. Andererseits kann unter der Voraussetzung des Archimedischen Axioms aus dem Cantorschen Axiom die Lückenlosigkeit gefolgert werden. Es kommt nun darauf an, zu erkennen, daß auch für nicht-archimedische Größensysteme und Körper das Cantorsche Axiom in der angegebenen Fassung erfüllt sein kann.

Daß der Nachweis hierfür nur durch Betrachtung komplizierterer Modelle möglich ist, beruht auf dem Umstand, daß unter einer gewissen ordnungstheoretischen Bedingung, welche eine Abschwächung des Archimedischen Postulates bildet, die Erfüllung des Cantorschen Axioms für einen Körper diejenige des Archimedischen Axioms zur Folge hat. Diese Bedingung lautet:

(3) *Jedes nicht leere und nach rechts offene Anfangsstück der Ordnung ist mit einer abzählbaren Folge konfinal,*

d.h. (nach Hausdorff), es gibt in dem Anfangsstück eine abzählbare Folge, welche zu jedem Element des Anfangsstückes ein späteres Glied enthält.

Daß in der Tat ein nicht-archimedischer Körper, welcher die Bedingung (3) erfüllt, nicht dem Cantorschen Axiom genügen kann, ergibt sich folgendermaßen. Sei a ein positives Element und b im Verhältnis zu a unendlich groß; dann bildet die Menge der Elemente, die im Verhältnis zu b unendlich klein sind, ein nicht leeres und nach rechts offenes Anfangsstück der Ordnung des Körpers. Auf Grund der Bedingung (3) ist dieses Anfangsstück mit einer abzählbaren (aufsteigenden) Folge $a_1, a_2, \ldots$ konfinal. Nehmen wir nun für b_n das Element $b/2^n$, so erkennt man sofort, daß die Folge der Intervalle (a_n, b_n) eine Verletzung des Cantorschen Axioms darstellt. (Die gleiche Überlegung ist auch allgemein auf ein Größensystem anwendbar, sofern dieses zu jedem Element auch die Hälfte enthält.)

Das gewünschte Modell eines nicht-archimedischen, aber dem Cantorschen Axiom genügenden Körpers muß hiernach gegen die Bedingung (3) verstoßen.

Zur Aufstellung eines solchen Modells können wir so verfahren[4]). Wir bilden zunächst ein geordnetes System S, welches die Vereinigung von

[4]) Die Methode der Konstruktion wurde von P. HERTZ gemeinsam mit dem Verfasser entwickelt. Vgl. P. HERTZ: Sur les axiomes d'Archimède et de Cantor. C. R. Soc. phys. et d'histoire nat. de Genève **51**, 179—181 (1934).

Systemen S_ν ist, wo ν eine Zahl der ersten oder zweiten Zahlenklasse ist und $S_\mu < S_\nu$ für $\mu < \nu$. S_0 besteht aus nur einem Element. $S_{\nu+1}$ wird aus S_ν durch folgende Hinzufügungen gewonnen: Für jedes Element, zu dem in S_ν kein unmittelbar folgendes bzw. vorhergehendes existiert, wird ein unmittelbar folgendes bzw. vorhergehendes hinzugefügt. Ferner wird zwischen zwei unmittelbar aufeinanderfolgende Elemente von S_ν je ein Element eingeschaltet. Zur Gewinnung von S_λ für eine Limeszahl λ bildet man zuerst $\sum_{\nu<\lambda} S_\nu$, fügt dann an sämtlichen Lückenstellen der erhaltenen Ordnung ausfüllende Elemente ein und fügt ferner noch ein erstes und ein letztes Element hinzu.

Das so gewonnene geordnete System S hat insbesondere die folgenden Eigenschaften: (a) Die Ordnung ist dicht; (b) das Cantorsche Axiom ist erfüllt; (c) zu jeder aufsteigenden Folge von Elementen, die einem Element b vorausgehen, gibt es ein Element vor b, dem auch noch alle Elemente der Folge vorausgehen, und das gleiche gilt auch für die umgekehrte Ordnung.

Aus dem System S erhalten wir eine additive Gruppe A durch Bildung von endlichen oder abzählbar wohlgeordneten Summen $\sum_{\varkappa<\nu} c_\varkappa s_\varkappa$, worin die $s_\varkappa$ (die Komponenten) Elemente aus S sind, deren Aufeinanderfolge in der Summe ihrer Ordnung in S entspricht, und die $c_\varkappa$ (die Koeffizienten) reelle Zahlen $\neq 0$ sind. Auch die leere Summe ist zugelassen; sie bildet das 0-Element von A. Als Vorzeichen einer nicht leeren Summe gilt das Vorzeichen von c_0. Zwei Elemente von A lassen sich komponentenweise addieren und subtrahieren, da ja die Vereinigung zweier entsprechend der Ordnung in S wohlgeordneter, höchstens abzählbarer Mengen von Elementen aus S wieder eine solche Menge liefert. Ein Element a von A heißt dann und nur dann kleiner bzw. größer als das Element b, wenn die Differenz $b-a$ positiv bzw. negativ ist. Gemäß dieser Festsetzung ist von zwei Elementen aus A mit verschiedenen Anfangskomponenten dasjenige, dessen Anfangskomponente in der Ordnung von S vorausgeht, in A das größere und sogar unendlich groß im Verhältnis zu dem anderen. Ferner besitzt die Ordnung von A die vorhin für S festgestellten Eigenschaften (a), (b) und (c). Die Dichtigkeit ersieht man ohne weiteres; die Eigenschaft (c) ergibt sich aus dem Umstande, daß zu jeder absteigenden Folge $a_1, a_2, \ldots$ von positiven Elementen aus A stets noch ein kleineres positives Element existiert; für die Verifikation der Erfüllung des Cantorschen Axioms muß man die verschiedenen möglichen Fälle diskutieren, wobei die Eigenschaften (b) und (c) der Ordnung von S, sowie die Lückenlosigkeit des Kontinuums wesentlich benutzt werden.

Ähnlich wie beim Übergang von S zu der additiven Gruppe A gehen wir nun von A zu einem Ring B über: Wir bilden wiederum endliche und abzählbar wohlgeordnete Summenausdrücke, aber jetzt mit Verwendung einer Unbestimmten t: $\sum_{\varkappa<\nu} c_\varkappa t^{a_\varkappa}$, worin die Koeffizienten $c_\varkappa$ reelle Zahlen $\neq 0$ sind; die Exponenten sind Elemente von A, und zwar sind sie in der Summe ansteigend wohlgeordnet. Die Addition und Subtraktion der Elemente von B

erfolgt ganz entsprechend wie bei den Elementen von A. Die leere Summe bildet wiederum das 0-Element. Auch die Bestimmung des Vorzeichens und die Definition von Kleiner und Größer ist entsprechend wie in A. Damit ergibt sich auch ohne weiteres, daß die Eigenschaften (a), (b) und (c) für die Ordnung von B zutreffen.

Die Multiplikation zweier Ausdrücke

$$\sum_{\varkappa<\nu} c_\varkappa t^{a_\varkappa}, \qquad \sum_{\mu<\nu'} c'_\mu t^{a'_\mu}$$

erfolgt distributiv, wobei zur Bildung des Produktes

$$c_\varkappa \cdot c'_\mu t^{a_\varkappa + a'_\mu}$$

die Addition in der Gruppe A verwendet wird. Zum Nachweis, daß auf diese Weise wieder ein Ausdruck aus B resultiert, ist zweierlei zu zeigen: (1) daß jeder Wert einer Summe $a_\varkappa + a'_\mu$ nur endlich oft auftritt, (2) daß die vorkommenden Summenwerte in der Ordnung nach der Größe wohlgeordnet sind. Um das erste zu beweisen, betrachten wir die Summen $a_\varkappa + a'_\mu$, welche einen gemeinsamen Wert a haben. Die zugehörigen Summanden $a_\varkappa$ sind aufsteigend wohlgeordnet. Demgemäß sind die Summanden a'_μ $(= a - a_\varkappa)$ absteigend wohlgeordnet. Andererseits aber sind sie auch (durch die Indizes) aufsteigend wohlgeordnet, also bilden sie eine endliche Gesamtheit. (2) ergibt sich ohne weiteres daraus, daß jede Teilmenge der Summenwerte $a_\varkappa + a'_\mu$ ein kleinstes Element hat.

Die multiplikative Einheit in B ist die eingliedrige Summe mit dem Koeffizienten 1 und dem 0-Element von A als Exponenten. Ferner ist ersichtlich, daß das Produkt zweier positiver Elemente wieder positiv, und das Produkt aus einem positiven und einem negativen Element negativ ist. Es gibt somit keinen Nullteiler in dem Ring B. Hiernach ist es möglich, von B zu dem Quotientenkörper Q überzugehen.

Auf den Körper Q läßt sich die Ordnung von B mittels der Vorzeichenbestimmungen in B ausdehnen. Für die so resultierende Ordnung gilt wegen der Beschaffenheit der Ordnung von B, daß zwischen je zwei verschiedenen Elementen von Q ein Element von B liegt. (Wie im §5 ausgeführt wird, fällt sogar der Körper Q mit B zusammen.) Damit ergibt sich, daß die Ordnung von Q dicht ist und ferner auch dem Cantorschen Axiom genügt, da dieses ja durch B erfüllt wird. Offensichtlich ist Q ein nicht-archimedischer Körper. Somit ist das gewünschte Modell gewonnen. —

Durch die am Anfang dieses Paragraphen angestellten Überlegungen werden wir auf die Möglichkeit geführt, *das System der reellen Zahlen zu charakterisieren, als einen solchen geordneten Körper, dessen Ordnung das Cantorsche Stetigkeitsaxiom erfüllt und außerdem die Bedingung, daß jedes nicht leere und nach rechts offene Anfangsstück mit einer abzählbaren Folge konfinal ist.*

Die gleiche Kennzeichnung der Stetigkeit ist auch für ein Größensystem möglich, wenn noch die Bedingung der Halbierbarkeit jeder Größe hinzugenommen wird. Andererseits kann hier auch beidemal anstatt der Bedingung (3) das Archimedische Axiom beibehalten werden. In diesem Falle

genügt bei den Größensystemen an Stelle der Bedingung der Halbierbarkeit die Forderung, daß jede Größe in kleinere zerlegbar ist.

Im Fall der Beibehaltung des Archimedischen Axioms kann statt der rein ordnungstheoretischen Fassung des Cantorschen Axioms auch die metrische Fassung genommen werden, bei welcher für die Intervallfolge die Bedingung hinzutritt, daß die Intervall-Längen unter jede Größe sinken[5]). Dabei ist es im Sinne der Absonderung des Cantorschen Axioms von dem archimedischen Axiom wünschenswert, diese Größenbedingung so zu fassen, daß nicht beliebig kleine Bruchteile der Einheit, sondern beliebig kleine Größen des betreffenden Größensystems gemeint sind. Das Cantorsche Axiom in dieser Fassung ist, wie man leicht nachweist, gleichbedeutend mit derjenigen Vollständigkeitsbedingung, welche besagt, daß jede dem Cauchyschen Konvergenzkriterium (im Sinne des genannten Größenbegriffes) genügende Folge einen Limes hat[6]). Die Unabhängigkeit des Archimedischen Axioms von dem so gefaßten Cantorschen Axiom ist aus einem von Hilbert gegebenen Modell zu entnehmen[7]).

§ 5. Der bei der Modellkonstruktion in § 4 verwendete Übergang von dem Ring B zu dem Quotientenkörper Q läßt sich als entbehrlich erkennen, wofür allerdings eine eingehendere Betrachtung erforderlich wird.

Nämlich wir können zeigen, daß der Ring B selbst schon ein Körper ist. Der Nachweis dafür besteht darin, festzustellen, daß zu jedem von 0 verschiedenen Element b aus B ein reziprokes existiert. Ein solches Element b ist darstellbar in der Form

$$c_0 t^{a_0} \cdot \left(1 + \frac{c_1}{c_0} t^{a_1 - a_0} + \cdots\right), \qquad c_0 \neq 0,$$

wobei der Klammerausdruck die Form $1+u$ hat und in u alle auftretenden Exponenten positive Elemente von A sind. Nun können wir das Reziproke von b ansetzen in der Form

$$\frac{1}{c_0} t^{-a_0}(1 - u + u^2 - u^3 + \cdots).$$

Daß die Entwicklung dieses Ausdruckes nach Potenzen von t möglich ist und ein zu b reziprokes Element liefert, wird erwiesen, indem wir zeigen, daß die Exponenten von t, die in den Entwicklungen der Potenzen u^k $(k = 1, 2, \ldots)$ auftreten, sich aufsteigend wohlordnen lassen und daß ferner

[5]) Hinsichtlich der Beziehungen zwischen diesen beiden Fassungen des Cantorschen Axioms vgl. die Abhandlungen von R. Baldus: Zur Axiomatik der Geometrie III, Sitzgsber. Heidelbg. Akad. Wiss., Math.-nat. Klasse **1930**, 5. Abh., und von A. Schmidt: Die Stetigkeit in der absoluten Geometrie, Sitzgsber. Heidelbg. Akad. Wiss., Math.-nat. Klasse **1931**, 5. Abh.

[6]) Auf diese Gleichwertigkeit wurde ich seinerzeit von R. Sikorski hingewiesen. Es sei bei dieser Gelegenheit auf die interessanten Beispiele von nicht-archimedischen Körpern aufmerksam gemacht, die Sikorski aufgestellt hat: On an ordered algebraic field. C. R. Soc. Sci. et des Lettres de Varsovie **41**, Cl. III, 69—96 (1948).

[7]) Vgl. die in Fußnote [5]) genannte Abhandlung von A. Schmidt, S. 6.

für jeden solchen Exponenten e nur endlich viele Koeffizienten von t^e zu addieren sind.

Dazu überlegen wir zunächst folgendes: Das Element u von B hat die Form $\sum_{\varkappa<\nu} q_\varkappa t^{d_\varkappa}$ wobei die $d_\varkappa$ positiv und aufsteigend wohlgeordnet sind. Die Potenz u^k hat demnach die Gestalt einer Summe von Gliedern $c\,t^e$, worin die Exponenten aufsteigend wohlgeordnet sind und jeder von ihnen eine k-gliedrige Summe von Exponenten $d_\varkappa$ ist. Hiernach besteht die Aufgabe darin, zu zeigen: (1) Die Werte der endlichen Summen aus den $d_\varkappa$ sind aufsteigend wohlgeordnet. (2) Jeder Summenwert wird nur durch endlich viele verschiedene Summen aus den $d_\varkappa$ dargestellt.

Sei M die Menge dieser Summenwerte; M ist eine Teilmenge von A. Die Ordnung der Elemente von M nach wachsender Größe bestimmt sich in erster Linie nach der Anfangskomponente, in zweiter Linie nach dem Koeffizienten der Anfangskomponente und dann nach den weiteren Gliedern. Die Anfangskomponente eines Elementes e aus M ist identisch mit der Anfangskomponente des größten Summanden in einer jeden der Summendarstellungen von e. Sie stimmt also überein mit der Anfangskomponente eines $d_\varkappa$. Diese Anfangskomponenten sind aber, so wie die $d_\varkappa$ selbst, aufsteigend wohlgeordnet.

Zum Nachweis von (1) genügt es demnach zu zeigen, daß die Elemente von M mit fester Anfangskomponente s wohlgeordnet sind, wobei im Sinne transfiniter Induktion vorausgesetzt werden kann, daß dieses für kleinere vorkommende Anfangskomponenten zutrifft.

Das geschieht nun auf folgende Weise. Jede Summe von Exponenten $d_\varkappa$, welche ein Element aus M mit der Anfangskomponente s darstellt, besteht zum Teil aus Gliedern mit der Anfangskomponente s, welche den „Hauptteil" der Summe bilden, und einem „Rest". In dem Hauptteil hat s einen bestimmten Koeffizienten, der zugleich der Koeffizient von s in der ganzen Summe ist. Sei c einer der so vorkommenden Koeffizienten, und sei $M_{s,c}$ die Menge der Elemente aus M mit der Anfangskomponente s und einem Koeffizienten $\leq c$. In der Wohlordnung der $d_\varkappa$ nach aufsteigender Größe gibt es ein erstes solches mit der Anfangskomponente s, wobei c_0 der Koeffizient von s sei. Es können nun die Hauptteile von darstellenden Summen für Elemente von $M_{s,c}$ aus höchstens $\left[\frac{c}{c_0}\right]$ Summanden bestehen. Diese höchstens $\left[\frac{c}{c_0}\right]$-gliedrigen Summen sind aber aufsteigend wohlgeordnet, wie es ja schon bei der Bildung der Potenzen u^k benutzt wird. Somit sind die Hauptteile in den Summendarstellungen der Elemente von $M_{s,c}$ aufsteigend wohlgeordnet. Zugleich sind aber die Reste in diesen Summendarstellungen auf Grund unserer Induktionsannahme wohlgeordnet; folglich sind auch die Elemente von $M_{s,c}$ aufsteigend wohlgeordnet und somit auch die Elemente von M mit der Hauptkomponente s. Damit ist die Behauptung (1) bewiesen.

Zum Nachweis von (2) machen wir wiederum eine transfinite Induktion nach der Anfangskomponente s des betreffenden Elementes e von M. c sei der Koeffizient von s in e. Auf Grund von (1) sind die Hauptteile der

darstellenden Summen von e aufsteigend wohlgeordnet; daher sind die Reste in diesen Summen absteigend wohlgeordnet. Andererseits sind diese, wiederum nach (1), auch aufsteigend wohlgeordnet, also haben wir nur endlich viele Werte von Hauptteilen und von Resten. Für jeden Wert des Hauptteiles gibt es aber wegen der durch c beschränkten Gliederzahl nur endlich viele Summendarstellungen, und für die Werte der Reste folgt das gleiche aus unserer Induktionsannahme. Somit haben wir nur endlich viele Summendarstellungen von e.

Hiermit sind nun die Behauptungen (1) und (2) über die Menge M als zutreffend erkannt. Daraus folgt zunächst die Möglichkeit, die unendliche Summe $1-u+u^2-u^3+\cdots$ in B zu bilden. Gleichermaßen ergibt sich allgemein, daß für ein Element u von B, worin alle Exponenten von t positiv sind (d.h. ein gegenüber 1 unendlich kleines Element), eine beliebige Potenzreihe

$$r_0+r_1u+r_2u^2+\cdots$$

mit reellen Koeffizienten $r_0, r_1, r_2 \ldots$ in B gebildet werden kann. Dabei entnehmen wir aus (2), daß in der Entwicklung des resultierenden Elementes von B nach Potenzen von t jeder vorkommende Exponent nur in der Entwicklung endlich vieler Potenzen von u, hernach aber nicht mehr auftritt. Dieser Umstand hat weiter auch zur Folge, daß die Multiplikation zweier aus Potenzreihen in u gewonnener Elemente so ausgeführt werden kann, daß zuerst die Reihen in u formal ausmultipliziert werden und hernach die Entwicklung nach Potenzen von t erfolgt. Es bleiben daher auch alle formal algebraischen Identitäten zwischen u-Reihen für die entsprechenden Elemente aus B gültig.

Insbesondere folgt hiernach, daß das Element von B, welches durch die Reihe $1-u+u^2-u^3+\cdots$ dargestellt wird, zu $1+u$ reziprok ist. Damit ist nun der Ring B als ein Körper erwiesen. Überdies ergeben sich aber noch bezüglich dieses Körpers etliche weitere Feststellungen. So insbesondere, daß für jede natürliche Zahl n zu jedem positiven Element von B eine n-te Wurzel existiert. Gleichermaßen folgt, daß im Bereich der gegenüber 1 unendlich kleinen Elemente die durch Potenzreihen definierten elementaren transzendenten Funktionen wie e^u, $\log(1+u)$, arc tg u usw. in B existieren, wobei die Additionstheoreme dieser Funktionen gültig bleiben.

Eidgenössische Technische Hochschule Zürich (Schweiz).

(Eingegangen am 16. März 1955.)

BOCHNER, S.
Math. Zeitschr. Bd. 63, S. 230—242 (1955)

Green-Goursat Theorem.

In memory of ISSAI SCHUR.

By

S. BOCHNER.

In a recent paper [1] we have generalized from complex to real variables GOURSAT's classical proof for the formula

$$\oint f(z)\,dz = 0 \tag{1}$$

in which only the existence of the derivative

$$\lim_{\zeta \to 0} \frac{f(z+\zeta) - f(z)}{\zeta} = f'(z) \tag{2}$$

is being assumed, but nothing further on the behaviour of $f'(z)$ as a function of the variable point z is being stipulated (explicitly).

In the present paper our previous theorem will be rounded off by incorporating two further generalizations, both noteworthy, we think (part I). Then, a pecularity of the one-dimensional case will be commented upon (part II), and finally an effect of the entire generalization on the CAUCHY formula with remainder term,

$$f(z,\bar{z}) = \frac{1}{2\pi i}\int_C \frac{f(\zeta,\bar{\zeta})\,d\zeta}{\zeta - z} - \frac{1}{2\pi i}\int_D \frac{\frac{\partial f(\zeta,\bar{\zeta})}{d\bar{\zeta}}\,d\bar{\zeta}\,d\zeta}{\zeta - z}, \tag{3}$$

will be discussed (part III).

Most modern papers on GREEN and STOKES theorem are very much more elaborate and refined than is the present one (see for instance [2] and authors there quoted), but we do not think that our theorems 1 and 2 are actually directly contained in any of them or that the direction or aim of the present paper has been anticipated as such.

Part I. Green's Theorem for Two Variables.

Theorem 1. *Let D be a bounded (open) domain in the (x, y)-plane and let C be its boundary. The formula*

$$\int_C A\,dy - B\,dx = \int_D (A_x + B_y)\,dx\,dy \tag{4}$$

holds under the following assumptions 1.—5.

1. *The boundary C consists of one or several simple closed rectifiable curves.*

2. *The functions $A(x, y)$, $B(x, y)$ are defined and continuous in the closure $\overline{D} = D + C$.*

3. *At each point x_0, y_0 of D the functions A, B have* STOLZ *derivatives, meaning that*

$$(5)\quad \begin{cases} A(x,y) = A(x_0,y_0) + A_{x_0}\cdot(x-x_0) + A_{y_0}\cdot(y-y_0) + o(|z-z_0|) \\ B(x,y) = B(x_0,y_0) + B_{x_0}\cdot(x-x_0) + B_{y_0}\cdot(y-y_0) + o(|z-z_0|), \end{cases}$$

where $z = x + iy$.

4. *The sum*

$$(6)\qquad \Phi(x,y) \equiv A_x + B_y$$

is LEBESGUE *integrable over D, meaning that*

$$(7)\qquad \int_D |A_x + B_y|\, dx\, dy < \infty.$$

5. *At each point x_0, y_0 of D we have the generalized continuity property*

$$(8)\qquad \lim_{\varepsilon \downarrow 0} \frac{1}{\pi\varepsilon^2} \int_{|z-z_0|<\varepsilon} |\Phi(x,y) - \Phi(x_0,y_0)|\, dx\, dy = 0.$$

Remark 1. In [*1*] we have established this theorem under the more special assumptions that $\Phi(x, y)$ is continuous (in the ordinary sense) and that the curves C are rectilinear polygons, and the present more general version would reflect the natural scope of the theorem, if it were not for the fact that the continuity property 5 can be relaxed to a certain extent, for a sufficiently smooth boundary at any rate (compare Remark 9 in part II). In this connection we would like to point out that according to a theorem of LEBESGUE,—for its one-dimensional version see [*3*], p. 13,—any LEBESGUE integrable function $\Phi(x, y)$ in D possesses the continuity property (8) at the points (x^0, y^0) of D automatically, but with the possible exception of a set of measure 0, whereas in our theorem no *general* set of measure 0 could be excepted, although certain null sets of special structure could.

Remark 2. As already emphasized in [*1*] the novelty of this version of GREEN's theorem lies in the fact that no assumption on the behaviour of any of the functions A_x, B_y and also A_y, B_x *separately* is being made, but only on their combination (6) which is the only one occurring explicitly in the formula.

Remark 3. If we introduce the differential operators

$$(9)\qquad \frac{\partial}{\partial z} = \frac{1}{2}\frac{\partial}{\partial x} + \frac{1}{2i}\frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\frac{\partial}{\partial x} - \frac{1}{2i}\frac{\partial}{\partial y},$$

and if we designate an arbitrary function $F(x, y)$ symbolically by $f(z, \bar{z})$ then our theorem implies

$$(10)\qquad \int_C f(z,\bar{z})\, dz = \int_D \frac{\partial f}{\partial \bar{z}}\, d\bar{z}\, dz, \qquad d\bar{z}\, dz = 2i\, dx\, dy,$$

under the assumption that f is continuous in $D + C$, and is STOLZ differentiable in D as a function of (x, y), and that the function $\frac{\partial f}{\partial \bar{z}}$ is LEBESGUE integrable over D and has the property (8) at each point of it.

Remark 4. It is easily verified that if $\Phi(x, y)$ has property (8) then for any continuous factor $\lambda(x, y)$ the product $\Psi(x, y) = \Phi(x, y)\,\lambda(x, y)$ will have it likewise. Also if D_ε is a neighborhood of (x_0, y_0) which contains a circle $|z - z_0| < \alpha\,\varepsilon$ and is itself contained in a circle $|z - z_0| < \beta\,\varepsilon$, where α, β are independent of ε, then (8) implies

$$\int_{D_\varepsilon} |\Psi(x, y) - \Psi(x_0, y_0)|\, dx\, dy = o(\varepsilon^2), \qquad \varepsilon \to 0. \tag{11}$$

Remark 5. If $x = x(u, v)$, $y = y(u, v)$ is a continuously differentiable one-one mapping of a domain S of the (u, v)-plane into our domain D, and if

$$J(u, v) \equiv \frac{\partial(x, y)}{\partial(u, v)} > 0, \tag{12}$$

then we have

$$\int_D \Phi(x, y)\, dx\, dy = \int_S \Psi(u, v)\, du\, dv, \tag{13}$$

where $\Psi(u, v) = \Phi(x, y)\, J(u, v)$, in the sense that if either integral (13) exists then so does the other. Furthermore if (u_0, v_0) maps into (x_0, y_0), then for small increments we have

$$\Delta z_0 = 0(\Delta w_0) \quad \text{and} \quad \Delta w_0 = 0(\Delta z_0),$$

and remark 4 leads to the following conclusion. If $\Phi(x, y)$ possesses property (8) as function in (x, y) then $\Psi(u, v)$ also possesses this property as function in (u, v).

Remark 6. If on a joint interval $0 \leq t \leq 1$ a sequence of functions $z_n(t) = x_n(t) + i\,y_n(t)$ is convergent to $z(t)$, if all $z_n(t)$ are of bounded variation and their variations are jointly bounded,

$$V_{0,1}\{z_n(\cdot)\} \leq M, \tag{14}$$

and if a sequence of continuous functions $\varphi_n(t)$ is uniformly convergent to $\varphi(t)$ on $0 \leq t \leq 1$, then

$$\lim_{n \to \infty} \int_0^1 \varphi_n(t)\, dz_n(t) = \int_0^1 \varphi(t)\, dz(t) \tag{15}$$

[RIEMANN-STIELTJES integrals, say]. This implies as follows.

Remark 7. If on a closed pointset $\overline{D}$ in (x, y)-plane there is given a sequence of (continuous) rectifiable curves $C_n: z_n(t)$, $0 \leq t \leq 1$; if they converge uniformly in $0 \leq t \leq 1$ to the curve $C: z(t)$; if the lengths of the curves C_n are jointly bounded; and if $A(x, y)$ and $B(x, y)$ are continuous functions on $\overline{D}$; then we have

$$\lim_{n \to \infty} \int_{C_n} A\big(x_n(t), y_n(t)\big)\, dy_n(t) - B\big(x_n(t), y_n(t)\big)\, dx_n(t)$$
$$= \int_C A\big(x(t), y(t)\big)\, dy(t) - B\big(x(t), y(t)\big)\, dx(t),$$

that is

$$\lim_{n \to \infty} \int_{C_n} A\, dy - B\, dx = \int_C A\, dy - B\, dx. \tag{16}$$

Remark 8. In the plane of $w=u+iv$ let S be a bounded domain which is bounded by the (outer) circle

$$R^0: w = w^0 + r^0 e^{i\vartheta}, \quad 0 \leq \vartheta < 2\pi$$

and (perhaps) by the additional (inner) circles

$$R^m: w = w^m + r^m e^{i\vartheta}, \quad m = 1, \ldots, n.$$

If for $0<\varepsilon<1$ we introduce the circles

$$R^0: w = w^0 + (1-\varepsilon) r^0 e^{i\vartheta}$$
$$R^m: w = w^m + (1+\varepsilon) r^m e^{i\vartheta}$$

then for ε sufficiently small they bound a domain S_ε interior to S which for $\varepsilon \to 0$ converges to the latter. Now, let $z=f(w)$ be a conformal mapping from the domain S into our previous domain D which also carries the circles R^0, R^m into rectifiable curves C^0, C^m respectively. It then carries the approximating circles R_ε^0, R_ε^m into approximating curves C_ε^0, C_ε^m which are uniformly convergent to the boundary curves C^0, C^m in parameter, and, what is decisive for us, the lengths of the curves C_ε^0, C_ε^m are jointly bounded in ε, see [4], p. 194. Thus, if we put $C_\varepsilon = C_\varepsilon^0 + C_\varepsilon^1 + \cdots + C_\varepsilon^m$, then by remark 7 we have

$$\lim_{\varepsilon \to 0} \int_{C_\varepsilon} A\,dy - B\,dx = \int_C A\,dy - B\,dx \tag{17}$$

for any functions A, B continuous in $D+C$.

Proof of Theorem 1. Following the classical procedure of GOURSAT we first assume, as we did in [1], that the domain D is a rectilinear triangle, which we will denote by D_0, and that the assumptions 2.—5. of the theorem are fulfilled in an open neighborhood U of the closed set $\bar{D}_0 = D_0 + C_0$. Assuming (4) not to hold, we write,

$$\left| \int_{C_0} A\,dy - B\,dx - \int_{D_0} \Phi(x,y)\,dx\,dy \right| = \alpha > 0, \tag{18}$$

and we subdivide D_0 into 4 rectilinear triangles $D_1^{(\nu)}$, $\nu = 1, \ldots, 4$, any two of which are congruent so that their sides are half as long as the corresponding ones in D_0 itself. Due to the "additivity" of relation (4), assumption (18) implies that we have

$$\left| \int_{C_1} A\,dy - B\,dx - \int_{D_1} \Phi(x,y)\,dx\,dy \right| \geq \frac{\alpha}{4}$$

for at least one of the triangles $D_1^{(\nu)}$, which we have chosen fixed and denoted by D_1. Continuing in this manner we obtain a sequence of triangles $D_0 > D_1 > D_2 > D_3 \ldots$ with the following properties. Firstly we have

$$\left| \int_{C_n} A\,dy - B\,dx - \int_{D_n} \Phi(x,y)\,dx\,dy \right| \geq \frac{\alpha}{4^n}; \tag{19}$$

secondly we have

$$\text{length } [C_n] = 0\left(\frac{1}{2^n}\right)$$

and $$\frac{1}{4^n} = 0\,(\text{area}\,[D_n])\,;$$

and thirdly we have that the closures $\overline{D}_n$ have a point in common, and this point we denote by x_0, y_0.

If now we put $A(x, y) = A' + A''$, $B = B' + B''$ where

$$A'(x, y) = A(x_0, y_0) + A_{x_0} \cdot (x - x_0) + A_{y_0} \cdot (y - y_0)$$
$$B'(x, y) = B(x_0, y_0) + B_{x_0} \cdot (x - x_0) + B_{y_0} \cdot (y - y_0),$$

and A'', B'' are corresponding "remainder functions", then the parts A', B' are linear in (x, y), and for such one we take our theorem as known, as does also GOURSAT in his case. However, since A, B are STOLZ differentiable at (x_0, y_0), we have

$$|A''| + |B''| = o\,(|z - z_0|) = o\left(\frac{1}{2^n}\right), \quad \text{for } z \text{ on } C_n;$$

and also

$$\Phi''(x, y) \equiv A''_x + B''_y = \Phi(x, y) - \Phi(x_0, y_0).$$

Therefore,

$$\int_{C_n} A''\,dy - B''\,dx = 0\,(\max_{z \in C_n} [|A''_n| + |B''_n|]) \cdot \text{length}\,[C_n] = o\left(\frac{1}{4^n}\right),$$

and

$$\int_{D_n} \Phi''\,dx\,dy = o\,(\text{area}\,[D_n]) = o\left(\frac{1}{4^n}\right), \tag{20}$$

and this would not be compatible with relation (19) for $A = A''$, $B = B''$; whence the conclusion.

The protective neighborhood U was needed because the point (x_0, y_0) in the proof might be on the boundary C of $D \equiv D_0$, but we can now dispense with it by the following reasoning. Let the origin $(0, 0)$ be a point of D and let $z = z(t)$ be a parametrization of C. If we introduce the triangle $C_\varepsilon: z_\varepsilon = (1 - \varepsilon)\,z(t)$, $0 < \varepsilon < 1$, then it bounds a subdomain D_ε of D of which D is such a neighborhood, so that our conclusion applies to D_ε at first. However, we have

$$\lim_{\varepsilon \to 0} \int_{D_\varepsilon} \Phi(x, y)\,dx\,dy = \int_D \Phi(x, y)\,dx\,dy$$

whenever the LEBESGUE integral on the right side exists, and on the other hand we have

$$\lim_{\varepsilon \to 0} \int_{C_\varepsilon} A\,dy - B\,dx = \int_C A\,dy - B\,dx$$

by remark 7, and thus (4) holds if D is a triangle.

Next, by combinatorial "additivity" it holds whenever the bounding curves $C^0, C^1, \ldots, C^n$ are rectilinear polygons. But then, again by the use of remark 7 ,it can be easily shown that it holds whenever the curves are arbitrarily convex, and especially whenever they are circles.

Turning now to the most general domain D as envisaged in the theorem, we introduce a conformal mapping of it into a domain S with circular boundary curves R, as in remark 8; and if, as in remark 8, we approximate to D by

subdomains D_ε whose boundaries C_ε map into circles R_ε of S, then on account of relation (17), our assertion will follow from the following lemma, finally.

Lemma 1. Let U be a neighborhood of $D+C$, and let $x=x(u,v)$, $y=y(u,v)$ be a one-one continuously differentiable mapping, with positive Jacobian, of the neighborhood U into a neighborhood V in (u,v)-plane, and let $S+R$ be the image of $D+C$, and let it be known that our theorem has already been established for this domain S; then it also holds for the domain D.

In fact, if we put

$$A_1(u,v) = A(x,y)\frac{\partial y}{\partial v} - B(x,y)\frac{\partial x}{\partial v}$$

$$B_1(u,v) = B(x,y)\frac{\partial x}{\partial u} - A(x,y)\frac{\partial y}{\partial u}$$

then we have identically

$$A\,dy - B\,dx = A_1\,du - B_1\,du$$

$$\left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y}\right)\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial A_1}{\partial u} + \frac{\partial B_1}{\partial v},$$

and all that is needed for the proof of the lemma has already been anticipated in remark 5.

Part II. Several Variables.

For more than two dimensions our theorem also holds (compare theorem 1) in [1]) but the case of "general" boundary surfaces is fraught with difficulties which we do not wish to enter into. However we do wish to make a statement resembling the present two-dimensional one, and we will make it for a certain type of simplex therefore.

For any $k \geq 1$, in the Euclidean $E_k:(u_j)$ we take a rectilinear k-dimensional simplex S, whose boundary we denote by R, and we take some open neighborhood V of $R+S$ also. We also take an arbitrary continuously differentiable mapping with positive Jacobian,

$$x_j = x_j(u_1, \ldots, u_k), \quad j = 1, \ldots, k$$

of the neighborhood V in $E_k:(u_j)$ into a neighborhood U in $E_k:(x_j)$, and the resulting image of (S,R) shall be denoted by (D,C) then.

Theorem 2. *In $E_k:(x_j)$ let D be any k-dimensional simplex as just described and let C be its boundary. Then the formula*

$$\int_C \sum_{j=1}^{k} A_j(x_k)\,(-1)^{j-1}\,dx_1 \ldots dx_{j-1}\,dx_{j+1} \ldots dx_k = \int_D \Phi(x_j)\,dv_x, \tag{21}$$

$$\Phi(x_j) \equiv \frac{\partial A_1}{\partial x_1} + \cdots + \frac{\partial A_k}{\partial x_k}, \tag{22}$$

holds under the following assumptions. Each of the functions

$$A_1(x_j), \ldots, A_k(x_j)$$

is defined and continuous in $D+C$ and has a STOLZ *derivative at each point (x_j^0) of D. The function $\Phi(x_j)$ is* LEBESGUE *integrable over D,*

$$\int_D |\Phi(x_j)|\, dv_x < \infty,$$

and at each point (x_j^0) of D we have

$$\int_{|x-x^0|<\varepsilon} |\Phi(x_j) - \Phi(x_j^0)|\, dv_x = o(\varepsilon^k), \qquad \varepsilon \to 0. \tag{23}$$

Remark 9. Assumption 5 in theorem 1 was needed only for the proof of the inequality (20), and therefore we may state as follows. If in theorem 1 (or 2) the domain D is a rectilinear polyhedron, then assumption (8), or (23), can be relaxed in the following manner. If $\{T\}$ is a family of k-dimensional rectilinear simplices in E_k any two of which are similar or antisimilar to one another and if each T contains the point (x_j^0) on its closure then we have

$$\int_T \left(\Phi(x_j) - \Phi(x_j^0)\right) dv_x = o(\text{volume } [T]), \tag{24}$$

as volume $[T]$ tends to 0, without absolute value signs under the integral.

Now, GOURSAT's method of proof for theorem 2 can also be made to apply to the one-dimensional case $k=1$, in which case D is an interval $a<x<b$, and C is the combinatorial pair of points $x=a$, $x=b$, and the resulting relation is

$$A(b) - A(a) = \int_a^b \frac{dA}{dx}\, dx. \tag{25}$$

Property (24) reads in this case

$$\lim_{h+k\to 0} \frac{1}{h+k} \int_{x^0-k}^{x^0+h} \frac{dA}{dx}\, dx = 0, \qquad h \geq 0,\ k \geq 0,\ h+k>0,$$

and thus the following statement is included in theorem 2.

Theorem 3. *Relation* (25) *holds whenever $A(x)$ is continuous in $a \leq x \leq b$, differentiable at each point of $a<x<b$, and $\frac{dA}{dx}$ is* LEBESGUE *integrable over $a<x<b$, and the function*

$$B(x) = \int_a^x \frac{dA}{d\xi}\, d\xi$$

has a derivative

$$\frac{dB(x)}{dx} \equiv \frac{dA(x)}{dx} \tag{26}$$

at each point in $a<x<b$.

This is a simple enough theorem, but it is not entirely "trivial" nevertheless. In fact, if we introduce the difference $F(x) = B(x) - A(x)$ then (26) implies

$$\frac{dF(x)}{dx} = 0, \tag{27}$$

and in order to conclude (25) it is necessary to know that (27) implies $F(x) = \text{constant}$, which is ROLLE's theorem, no less. Thus we see that, in the one-dimensional case, GOURSAT's method of proving (25) by subdividing the interval (a, b) into equal halves successively is as good as ROLLE's reasoning, and for $k \geq 2$ it is a rather efficient replacement for it.

We are now returning to the k-dimensional case, and to the following compare [5] and chapter VIII in [6].

We take a domain D and a subdomain D_0 whose closure $\overline{D}_0$ is contained in D, and we denote by ϱ the distance from D_0 to the boundary of D. If for $\sigma < \varrho$ we take in the sphere

$$S_\sigma : (t_1^2 + \cdots + t_k^2)^{\frac{1}{2}} \leq \sigma \tag{28}$$

an integrable function $K(t) \equiv K(t_1, \ldots, t_k)$, for which

$$\int_{S_\sigma} K(t)\, dv_t = 1, \tag{29}$$

say, and if a function $A(x) \equiv A(x_1, \ldots, x_k)$ is defined and continuous in D, then for x in D_0, we can introduce the "average"

$$A^K(x) = \int_{S_\sigma} A(x - t)\, K(t)\, dv_t \tag{30}$$

which, incidentally, for $\sigma \to 0$, converges to $A(x)$ there. This function (30) is a "smoothing" of $A(x)$ itself in the following manner (among others). If we replenish the function $K(t)$ in (28) to a function in the entire E_k by putting it $= 0$ for $|t| > \sigma$, and if we also replenish $A(x)$ to a function in E_k by putting it $= 0$ for x in $E_k - D$ then we can write

$$A^K(x) = \int_{E_k} A(x - t)\, K(t)\, dv_t \tag{31}$$

and also

$$A^K(x) = \int_{E_k} A(t)\, K(x - t)\, dv_t. \tag{32}$$

Now, in the first place, (32) implies that if $K(t)$ as a function in E_k belongs to differentiability class $C^{(1)}$ there, then $A^K(x)$ belongs to $C^{(1)}$ in E_k, and especially in D; and, in the second place, if $A(x)$ itself belongs to $C^{(1)}$ in D, then $A^K(x)$ belongs to $C^{(1)}$ in D_0 and we can also interchange differentiation with integration, thus

$$\frac{dA^K(x)}{\partial x_j} = \int_{S_\sigma} \frac{\partial A(x - t)}{\partial x_j} K(t)\, dv. \tag{33}$$

Now, if $A(x)$ has only a STOLZ derivative at each point then the second property cannot be asserted, but our theorem 2 will permit us to make a certain statement nevertheless.

Before making it, we define as follows. If $K(t)$ is integrable in (28) then we say that it is a $C^{(1)}$-kernel, if for every function $A(x)$ which is continuous

in D, the corresponding function (32) is continuous in D_0. As already remarked, if $K(t)$, as a function in E_k is $C^{(1)}$ there, then it is a $C^{(1)}$-kernel in S_σ, but there are important $C^{(1)}$-kernels which are more general than that. For instance for $h \leq \frac{\sigma}{k}$, the function

$$K(t) = \begin{cases} \frac{1}{(2h)^k}, & |t_j| \leq h,\ j = 1, \ldots, k \\ 0, & \text{some } |t_j| > 0, \end{cases}$$

is a $C^{(1)}$-kernel, the corresponding function (32) being the familiar average

$$\frac{1}{(2h)^k} \int_{-h}^{h} \cdots \int_{-h}^{h} A(x_1 + t_1, \ldots, x_k + t_k)\, dv_t$$

then.

Theorem 4. *If in a domain D in E_k we are given functions $A_1(x), \ldots, A_k(x)$, $\Phi(x)$ with exactly the same properties as in theorem 2, thus being connected by* (22) *among others, and if for a $C^{(1)}$-kernel $K(t)$ in S_σ we introduce the averages*

$$A_j^K(x) = \int_{S_\sigma} A_j(x - t)\, K(t)\, dv_t \tag{34}$$

$$\Phi^K(x) = \int_{S_\sigma} \Phi(x - t)\, K(t)\, dv_t \tag{35}$$

then we have the relation

$$\sum_{j=1}^{k} \frac{\partial A_j^K(x)}{\partial x_j} = \Phi^K(x) \tag{36}$$

for almost all points x in D_0. (If $\Phi^K(x)$ is continuous in D_0, then this holds at all points without exception.)

Proof. For x in D_0 and t in S_σ, the original relation (22) implies

$$\sum_{j=1}^{k} \frac{\partial A_j(x - t)}{\partial x_j} = \Phi(x - t)$$

and therefore

$$\int_{S_\sigma} \left(\sum_{j=1}^{k} \frac{\partial A_j(x - t)}{\partial x_j} \right) K(t)\, dv_t = \int_{S_\sigma} \Phi(x - t)\, K(t)\, dv_t,$$

and our theorem asserts, that under the assumptions of theorem 2, the left side can be re-arranged thus:

$$\sum_{j=1}^{k} \frac{\partial}{\partial x_j} \int_{S_\sigma} A_j(x - t)\, K(t)\, dv_t,$$

this involving two interlocking interchanges between summation, integration and differentiation.

We take in D_0 a function $\lambda(x) = \lambda(x_1, \ldots, x_k)$ which is $C^{(1)}$ there and which vanishes outside a polyhedral neighborhood U in D_0 (a so-called "testing

function" in the terminology of [5]). For fixed t in S_σ, the function in x,

$$\bar{A}_j(x) = \lambda(x)\, A_j(x-t); \tag{37}$$

has a STOLZ derivative at each point of D_0, and the "divergence"

$$\bar{\Phi}(x) = \sum_{j=1}^{k} \frac{\partial \bar{A}_j(x)}{\partial x_j} \equiv \lambda(x)\, \Phi(x-t) + \sum_{j=1}^{k} \frac{\partial \lambda(x)}{\partial x_j} A_j(x-t)$$

has the same properties in D_0, as $\Phi(x)$ has in D. However, all functions $\bar{A}_j(x)$ are 0 on the boundary of U, and thus by theorem 2 we have

$$\int_U \bar{\Phi}(x)\, dv_x = 0,$$

that is

$$\int_U \left[\lambda(x)\, \Phi(x-t) + \sum_{j=1}^{k} \frac{\partial \lambda}{\partial x_j} A_j(x-t) \right] dv_x = 0.$$

If we now multiply this by $K(t)$ and integrate with respect to t over S_σ, then by FUBINI's theorem the two integrations can be interchanged, and we have

$$\int_U \left[\lambda(x)\, \Phi^K(x) + \sum_{j=1}^{k} \frac{\partial \lambda}{\partial x_j} A_j^K(x) \right] dv_x = 0.$$

However, the functions $A_j^K(x)$ are $C^{(1)}$ and λ is 0 outside U, so that

$$\int_U \sum \frac{\partial \lambda}{\partial x_j} A_j^K\, dv_x = -\int_U \lambda \sum \frac{\partial A_j^K}{\partial x_j}\, dv_x,$$

and altogether we obtain

$$\int_{D_0} \lambda(x) \left[\Phi^K(x) - \sum_{j=1}^{k} \frac{\partial A_j^K}{\partial x_j} \right] dv_x = 0.$$

But, if this is so for every testing function $\lambda(x)$ as introduced, then we indeed must have relation (36) for almost all points of D_0, as claimed.

Part III. Cauchy Formula with Remainder Term.

In $E_2:(\xi, \eta)$ if $H(\xi, \eta)$ is LEBESGUE measurable and

$$\int_{E_2} \frac{|H(\xi, \eta)|}{1 + |\xi| + |\eta|}\, d\xi\, d\eta < \infty \tag{38}$$

then for almost all (x, y) the integral

$$G(x, y) = \int_{E_2} \frac{H(\xi, \eta)}{\zeta - z}\, d\xi\, d\eta \tag{39}$$

exists, $\zeta=\xi+i\eta$, $z=x+iy$, and $G(x,y)$ is LEBESGUE integrable over any bounded domain. This follows from FUBINI's theorem by showing that

$$\int_{|z|\leq A} dx\,dy \int_{E_2} \left|\frac{H(\xi,\eta)}{\zeta-z}\right| d\xi\,d\eta<\infty, \tag{40}$$

for any $A<\infty$ (integrate with respect to x, y first), and it is also easily seen that if $H(\xi,\eta)$ is bounded in a neighborhood, then the integral (39) exists at all points of the neighborhood without exception and is likewise bounded there.

Also, if we apply to the integral

$$\int_{|\zeta|\leq 1} \frac{H(\xi,\eta)\,d\xi\,d\eta}{\zeta-z}, \qquad |z|<1,$$

E. E. LEVI's transformation of coordinates

$$\xi-x=\varrho(\cos\vartheta-x), \qquad \eta-y=\varrho(\sin\vartheta-y)$$
$$0\leq\vartheta<2\pi, \qquad 0<\varrho\leq 1,$$

then it goes over into

$$\int_{\varrho=0}^{1}\int_{\vartheta=0}^{2\pi} h\left(x(1-\varrho)+\varrho\cos\vartheta,\ y(1-\varrho)+\varrho\sin\vartheta\right)\frac{1-\cos\vartheta\,x-\sin\vartheta\,y}{e^{i\vartheta}-z}\,d\varrho\,d\vartheta,$$

and we hence conclude that if the function $H(\xi,\eta)$ is $C^{(r)}$, $r\geq 1$, in a neighborhood U, then $G(x,y)$ is also $C^{(r)}$ there, and if $H(\xi,\eta)$ is analytic in the real variables ξ,η there, then $G(x,y)$ is likewise so in x, y. However, $G(x,y)$ cannot be a holomorphic function in z unless $H(\xi,\eta)$ is $\equiv 0$ in U, as the following theorems will imply.

Now, if in (10) we replace the letter z by ζ, and subsequently the function $f(\zeta,\bar{\zeta})$ by the function

$$\frac{f(\zeta,\bar{\zeta})}{\zeta-z},$$

then by the customary procedure of excising a small neighborhood around z the following theorem ensues.

Theorem 5. *If D and C are as in theorem 1, if $f(\zeta,\bar{\zeta})$ is continuous in $D+C$ and* STOLZ *differentiable in D, and if $f_{\bar{\zeta}}(\zeta,\bar{\zeta})$ is* LEBESGUE *integrable over D and possesses the continuity property 5 of theorem 1 there, then we have*

$$f(z,\bar{z})=\frac{1}{2\pi i}\int_C \frac{f(\zeta,\bar{\zeta})\,d\zeta}{\zeta-z}-\frac{1}{2\pi i}\int_D \frac{f_{\bar{\zeta}}(\zeta,\bar{\zeta})\,d\bar{\zeta}\,d\zeta}{\zeta-z} \tag{41}$$

at all points z of D at which the second integral exists (the exceptional points having measure zero at most).

If $f_{\bar{\zeta}}$ is also locally bounded in D, then the formula is valid at all points, and if $f_{\bar{\zeta}}$ is $C^{(r)}$ or real-analytic, then $f(\zeta,\bar{\zeta})$ itself is likewise so.

Of interest is the following companion theorem.

Theorem 6. *If $h(\zeta, \bar{\zeta})$ is* LEBESGUE *integrable on D so that the function*

$$g(z, \bar{z}) = -\frac{1}{2\pi i} \int_D \frac{h(\zeta, \bar{\zeta})\, d\bar{\zeta}\, d\zeta}{\zeta - z} \tag{42}$$

exist almost everywhere and is LEBESGUE *integrable over D itself, then for any testing function $\lambda(z, \bar{z})$ in D we have*

$$\int_D \left(\frac{\partial \lambda}{\partial \bar{z}} \cdot g + \lambda \cdot h \right) d\bar{z}\, dz = 0, \tag{43}$$

and if furthermore it so happens that $g(z, \bar{z})$ has a STOLZ *derivative everywhere and $g_{\bar{z}}$ has the suitable properties, then*

$$\frac{\partial g}{\partial \bar{z}} = h(z, \bar{z}) \tag{44}$$

almost everywhere.

Proof. From (42) we obtain

$$\int_D \frac{\partial \lambda}{\partial \bar{z}}\, g\, d\bar{z}\, dz = -\int_D h(\zeta, \bar{\zeta}) \left[-\frac{1}{2\pi i} \int_D \frac{\frac{\partial \lambda}{\partial \bar{z}}\, d\bar{z}\, dz}{z - \zeta} \right] d\bar{\zeta}\, d\zeta.$$

However, if in (41) we replace $f(z, \bar{z})$ by $\lambda(z, \bar{z})$ and recall that λ is 0 on the boundary of D, then we see that the inner integral has the value $\lambda(\zeta, \bar{\zeta})$, which proves (43). Next, as in the proof to theorem 4, if g has a STOLZ derivative everywhere, and $g_{\bar{z}}$ has the suitable properties (for instance if it is continuous) then (43) implies

$$\int_D \lambda \left(h - \frac{\partial g}{\partial \bar{z}} \right) d\bar{z}\, dz = 0$$

and this implies (44) a.e.

Theorem 5 (but not theorem 6) has a natural extension to k complex variables (see [7], p. 660/1). If we put

$$\omega = d\bar{\zeta}_1 d\zeta_1 + \cdots + d\bar{\zeta}_k d\zeta_k \quad \text{(external form)}$$

and

$$V_j = \frac{\bar{\zeta}_j - \bar{z}_j}{\left(\sum_p |\zeta_p - z_p|^2 \right)^k}$$

then a generalization of (41) is

$$(2\pi i)^k f(z_1, \bar{z}_1, z_2, \bar{z}_2, \ldots, z_k, \bar{z}_k) = \int_C f \sum_j V_j\, d\zeta_j\, \omega^{k-1} - \frac{1}{k} \int_D \sum_j \frac{\partial f}{\partial \bar{\zeta}_j} V_j \omega^k, \tag{45}$$

and an extension of theorem 5 is as follows.

Theorem 7. *If D is a polyhedral domain, and the function f has a* STOLZ *derivative in D, and if each of the k functions*

$$\frac{\partial f}{\partial \bar{\zeta}_1}, \ldots, \frac{\partial f}{\partial \bar{\zeta}_k} \tag{46}$$

is LEBESGUE *integrable and possesses the continuity property of theorem 2, then* (45) *holds at all points at which the second integral in* (45) *exists (the exceptional points having measure 0 at most).*

If the functions (46) *are also locally bounded in D, then the formula is valid at all points, and if they are $C^{(r)}$ as real-analytic then f itself is likewise so.*

References.

[1] BOCHNER, S.: Functions in one complex variable as viewed from the theory of functions in several variables, in "Lectures on Functions of a Complex Variable", p. 315. University of Michigan Press 1955. — [2] KRICKEBERG, K.: Über den Gaußschen und den Stokesschen Integralsatz, I and II. Math. Nachr. **10**, 261—314 (1953); **11**, 35—60 (1954). — [3] LEBESGUE, H.: Séries trigonométriques. 1906. — [4] SEIDEL, W.: Ränderzuordnung bei konformen Abbildungen. Math. Ann. **104**, 182—243 (1931). — [5] BOCHNER, S.: Linear partial differential equations with constant coefficients. Ann. of Math. **47**, 202—212 (1946). — [6] BOCHNER, S., and W. T. MARTIN: Several complex variables. Princeton 1948. — [7] BOCHNER, S.: Analytic and meromorphic continuation by means of Green's formula. Ann. of Math. **44**, 652—674 (1943).

Department of Mathematics, Princeton University, Princeton N. J. (USA.).

(Eingegangen am 2. April 1955.)

JACOBSTHAL, E.
Math. Zeitschr. Bd. 63, S. 243—276 (1955)

Über vertauschbare Polynome.

Dem Andenken an meinen verehrten Lehrer und Freund I. SCHUR gewidmet.

Von
ERNST JACOBSTHAL.

Einleitung.

Sind $f(x)$ und $g(x)$ zwei Polynome der Grade m und n, so sind die beiden Polynome $f(g(x))$ und $g(f(x))$ beide vom Grade $m \cdot n$, aber im allgemeinen verschieden voneinander. $f(x)$ und $g(x)$ heißen *vertauschbare*[1]) Polynome, wenn

$$f(g(x)) = g(f(x)) \tag{1}$$

ist. Für das zusammengesetzte Polynom $f(g(x))$ schreiben wir abgekürzt fg. Bei der Zusammensetzung kann eine der Komponenten oder beide den Grad 0 haben. Zu den Polynomen des Grades 0, die ja Konstanten sind, wollen wir hier auch die Zahl 0 rechnen. Für die Zusammensetzung gilt das assoziative Gesetz

$$(fg)h = f(gh). \tag{2}$$

Daher kann man die „Potenzen" $f^3 = fff$, f^4, ... definieren. Setzt man noch $ff = f^2$, $f^1 = f$, $f^0 = x$, so ist

$$f^n f^m = f^m f^n = f^{n+m}, \qquad n, m = 0, 1, 2, \ldots. \tag{3}$$

Die f^n sind die sog. „Iterierten" von f. Daß (3) auch im Falle $n = 0$ oder $m = 0$ gilt, beruht auf dem trivialen

Satz 1. *Jedes Polynom $f(x)$ ist mit $g(x) = x$ vertauschbar.*

Bei der Zusammensetzung spielt also $g(x) = x$ die Rolle der Einheit. — Die Veranlassung zu diesen Untersuchungen war eine spezielle Folge von Polynomen, die alle miteinander vertauschbar sind, und die Frage nach ihrer *Vollständigkeit*, d.h. die Frage: Gibt es außerhalb der Folge noch andere Polynome, die mit allen Elementen der Folge vertauschbar sind? Ist ein Polynom $g(x)$ gegeben, und fragt man nach allen Lösungen $f(x)$ der Gl. (1), so kennt

[1]) Über vertauschbare rationale Funktionen liegen funktionentheoretische Untersuchungen aus den zwanziger Jahren dieses Jahrhunderts vor (FATOU, JULIA, RITT). Die Literatur findet man zusammengestellt bei J. F. RITT: (1) Permutable Rational Functions. Trans. Amer. Math. Soc. **25**, 399—448 (1923). — (2) Prime and Composite Polynomials. Trans. Amer. Math. Soc. **23**, 51—66 (1922). — (3) JULIA, G.: Mémoire sur la permutabilité des fractions rationelles. Ann. Ecole norm. supérieure **39**, 131—215 (1922). — (4) FATON, P.: Sur un groupe de substitutions algébriques. C. R. Acad. Sci. Paris **173**, 694—696 (1921).

man ja gewisse Lösungen, nämlich die Iterierten $f(x)=g^n$ von $g(x)$. Es kann sein, daß es gar keine anderen Lösungen $f(x)$ gibt; es kann aber auch trivialerweise vorkommen, daß es andere Lösungen gibt, nämlich dann, wenn $g(x)$ selbst die Iterierte eines anderen Polynoms $h(x)$ ist. Aber selbst wenn $g(x)$ nicht die Iterierte eines anderen Polynoms $h(x)$ ist, kann es vorkommen, daß die Gl. (1) Lösungen $f(x)$ besitzt, die keine Iterierten von $g(x)$ sind. Bei den Fragen der Vertauschbarkeit nehmen die Polynome ersten und zweiten Grades eine Sonderstellung ein, wie wir noch sehen werden. Außerdem spielen die Lösungen der Schröderschen Gleichung $f(x)=x$ eine gewisse Rolle. Wir wollen sie die *Eigenwerte* des Polynomes $f(x)$ nennen.

§ 1. Die Eigenwerte eines Polynoms und seiner Iterierten.

Der Begriff des Eigenwerts eines Polynoms $y=f(x)$ ist identisch mit dem Begriff des Fixpunktes der Funktionentheorie. Das Polynom $f(x)\equiv x$ hat jede Zahl zum Eigenwert; das Polynom $f(x)=x+b$, $b\neq 0$, besitzt überhaupt keinen Eigenwert. In allen anderen Fällen besitzt ein Polynom positiven Grades n genau n Eigenwerte, die aber auch zusammenfallen können. Dann handelt es sich um mehrfache Eigenwerte.

Satz 2. *Die einzigen Konstanten, mit denen ein Polynom vertauschbar ist, sind seine Eigenwerte.*

Denn ist $f(x)$ mit einer Konstanten $g(x)=a$ vertauschbar, so ist $fg=f(a)$ und $gf=a$, also $f(a)=a$. Das Polynom $f(x)=x+b$, $b\neq 0$, ist mit keiner Konstanten vertauschbar.

Wir bezeichnen die aus $f^0=x, f^1, f^2, \ldots$ gebildete Folge mit V_f:

$$V_f=\{f^n\}_{n=0,1,2,\ldots}. \tag{4}$$

Ist der Grad von f größer als 1, so sind alle Elemente der Folge V_f voneinander verschieden. Ist der Grad m von $f(x)$ gleich 0, $f(x)$ eine Konstante b, so ist $f^1=f^2=\cdots$ und bei $f(x)\equiv x$ ist $f^0=f^1=f^2\equiv\cdots$. Im Falle $f(x)=x+b$, $b\neq 0$, ist $f^n=x+nb$, und alle Elemente in (4) sind voneinander verschieden. Ist schließlich $f(x)=ax+b$ mit $a\neq 1$ und $x_0=\frac{b}{1-a}$ der Eigenwert von $f(x)$, so ist $f(x)=a(x-x_0)+x_0$ und $f^n=a^n(x-x_0)+x_0$. Ist dann a keine Einheitswurzel, so sind alle Elemente in (4) voneinander verschieden und x_0 der einzige Eigenwert, den jedes f^n bei $n\geq 1$ hat. Ist aber $a\neq 1$ eine primitive r-te Einheitswurzel, so ist $f^n=f^m$ für $n\equiv m \bmod r$, die Folge (4) ist periodisch. Für $n\not\equiv 0 \bmod r$ hat f^n nur den Eigenwert x_0, und für $n\equiv 0 \bmod r$ ist $f^n\equiv x$ und hat jede Zahl zum Eigenwert.

Da für ein beliebiges Polynom $f(x)$, das ja abgesehen vom Falle $f(x)=x+b$, $b\neq 0$, stets einen Eigenwert hat, jeder Eigenwert auch Eigenwert von $f^n(x)$ ist, so verschwindet f^n-x für jede Nullstelle von $f(x)-x$. Daher ist zu erwarten, daß das Polynom $f^n(x)-x$ durch das Polynom $f(x)-x$ teilbar ist.

Man rechnet das so aus: Ist $f(x)=\sum_{\nu=0}^{m} a_\nu x^\nu$, so wird

$$f^2(x)-f(x)=\sum_{\nu=0}^{m} a_\nu\left[(f(x))^\nu-x^\nu\right]$$

$$=(f(x)-x)\sum_{\nu=0}^{m} a_\nu\frac{(f(x))^\nu-x^\nu}{f(x)-x}$$

$$=(f(x)-x)\,H_2(x),$$

worin $H_2(x)$ die letzte Summe ist und sich nach Ausrechnung gleich

$$H_2(x)=\sum_{\nu=0}^{m-1}(f(x))^\nu\sum_{\lambda=\nu+1}^{m} a_\lambda x^{\lambda-\nu-1} \tag{5}$$

ergibt. Also wird

$$f^2(x)-x=(f(x)-x)\,(1+H_2(x)). \tag{5a}$$

Setzt man nun allgemein

$$f^\nu(x)-x=(f(x)-x)\,(1+H_\nu(x)),\qquad H_0=-1,\ H_1=0, \tag{6}$$

so ist (6) für $\nu=0, 1, 2$ richtig. Indem man dann in der schon für den Index ν als richtig angenommenen Formel (6) x durch $f(x)$ ersetzt und (5) und (5a) benutzt, erhält man aus

$$f^{\nu+1}(x)-x=f^{\nu+1}(x)-f(x)+f(x)-x$$

nach kurzer Rechnung für die $H_\nu(x)$ die Rekursionsformel

$$H_{\nu+1}(x)=H_2(x)\,(1+H_\nu(f(x)))\qquad \nu\geq 0, \tag{7}$$

so daß (6) mit den so erhaltenen H_ν für alle $\nu\geq 0$ richtig ist.

Ist der Grad m des Polynomes $f(x)$ größer als 1, dann wird im allgemeinen schon $f^2(x)$ außer den Eigenwerten von $f(x)$ solche Eigenwerte besitzen, die keine von $f(x)$ sind. Das Beispiel $f(x)=x^2-\frac{3}{4}$ zeigt aber, daß es vorkommen kann, daß $f^2(x)$ keine neuen Eigenwerte aufweist. Für dieses Beispiel wird nämlich

$$f(x)-x=(x+\tfrac{1}{2})\,(x-\tfrac{3}{2}),$$

$$f^2(x)-x=(x+\tfrac{1}{2})^3(x-\tfrac{3}{2}).$$

Von den beiden einfachen Eigenwerten von $f(x)$ ist der eine ein einfacher, der andere ein dreifacher von $f^2(x)$ geworden. Die Bedingung dafür, daß ein Eigenwert ζ von $f(x)$ ein mehrfacher ist, lautet $f'(\zeta)=1$. Für $f'(\zeta)\neq 1$ ist ζ ein einfacher Eigenwert von $f(x)$. Für einen Eigenwert ζ von $f(x)$ ergibt (5) die Gleichung

$$H_2(\zeta)=f'(\zeta). \tag{5b}$$

Daher folgt aus (7) durch Rekursion:

$$H_\nu(\zeta)=f'(\zeta)+(f'(\zeta))^2+\cdots+(f'(\zeta))^{\nu-1},\qquad \nu\geq 2. \tag{7a}$$

Ist also ζ ein mehrfacher Eigenwert von $f(x)$, so ist wegen $f'(\zeta)=1$ nach (7a): $H_r(\zeta)=r-1$, also zeigt (6), daß jede der Iterierten f^r diesen Eigenwert genau von derselben Ordnung hat wie f selbst. Nur für einen einfachen Eigenwert ζ von $f(x)$, $f'(\zeta)\neq 1$, kann es vorkommen, daß er ein mehrfacher Eigenwert einer Iterierten f^r wird. Damit sich dies ereignet, muß nach (6) der zweite Faktor rechts für $x=\zeta$ verschwinden, also $H_r(\zeta)=-1$ sein. Da hier $f'(\zeta)\neq 1$ ist, folgt aus $H_r(\zeta)=-1$ nach (7a), daß

$$(f'(\zeta))^r=1, \quad f'(\zeta)\neq 1. \tag{7b}$$

Ist also für einen einfachen Eigenwert ζ von $f(x)$, d.h. für $f'(\zeta)\neq 1$, $f'(\zeta)$ eine r-te Einheitswurzel, so wird für die r-te Iterierte $f^r(x)$ dieses ζ ein mehrfacher Eigenwert. Dieser Eigenwert wird sogar ein mindestens dreifacher Eigenwert von $f^r(x)$. Ich will auf diesen letzten Punkt mit Rücksicht auf das oben gegebene Beispiel besonders hinweisen. Ich gebe den Beweis, ohne die Rechnungen durchzuführen. Es wird vorausgesetzt: ζ ein Eigenwert von $f(x)$, $f'(\zeta)\neq 1$, $(f'(\zeta))^r=1$. Man differenziert erst (5) und setzt dann $x=\zeta$. Das ergibt

$$H_2'(\zeta)=\tfrac{1}{2}f''(\zeta)\,(1+f'(\zeta)). \tag{a}$$

Dann differenziert man die Rekursionsformel (7) und setzt auch hierin $x=\zeta$. Mit Benutzung von (7a) erhält man dann, wenn man in (7) r durch λ ersetzt,

$$H_{\lambda+1}'(\zeta)=\frac{1}{2}f''(\zeta)\frac{f'(\zeta)+1}{f'(\zeta)-1}\left((f'(\zeta))^\lambda-1\right)+(f'(\zeta))^2H_\lambda'(\zeta), \quad \lambda=0,1,\ldots. \tag{b}$$

Setzt man nun

$$\frac{H_\lambda'(\zeta)}{(f'(\zeta))^{2\lambda}}=A_\lambda, \tag{c}$$

so geht (b) über in

$$A_{\lambda+1}=\frac{1}{2}f''(\zeta)\frac{f'(\zeta)+1}{f'(\zeta)-1}\left(\frac{1}{(f'(\zeta))^{\lambda+2}}-\frac{1}{(f'(\xi))^{2\lambda+2}}\right)+A_\lambda; \quad A_0=0. \tag{d}$$

Hieraus folgt durch Summation über $\lambda=0,1,\ldots,r-1$

$$A_r=\frac{1}{2}f''(\zeta)\frac{f'(\zeta)+1}{f'(\zeta)-1}\cdot\frac{1}{(f'(\zeta))^2}\cdot B_r, \tag{e}$$

worin

$$B_r=\sum_{\lambda=0}^{r-1}(f'(\zeta))^{r-\lambda}-\sum_{\lambda=0}^{r-1}(f'(\zeta))^{2(r-\lambda)}$$

ist. Für $f'(\zeta)=-1$ folgt aus (e) sofort, daß $A_r=0$ ist und für $f'(\zeta)\neq -1$ wird $(f'(\zeta))^2\neq 1$, und dann wird $B_r=0$, also auch $A_r=0$; also ergibt sich $H_r'(\zeta)=0$, und das zeigt, daß der letzte Faktor rechts in (6) ζ noch als mindestens zweifache Nullstelle besitzt, also ζ ein mindestens dreifacher Eigenwert von $f^r(x)$ ist. Wir haben also folgenden

Satz 3. *Ist $f(x)$ ein Polynom eines Grades ≥ 2, so ist jeder mehrfache Eigenwert von $f(x)$ für jedes $f^r(x)$, $r\geq 1$, ein Eigenwert von der gleichen Vielfachheit wie für $f(x)$. Ein einfacher Eigenwert ζ von $f(x)$, für den also $f'(\zeta)\neq 1$*

ist, ist dann und nur dann für eine Iterierte $f^r(x)$ ein mehrfacher Eigenwert, wenn $(f'(\zeta))^r = 1$ ist. Er ist dann sogar ein mindestens dreifacher Eigenwert von $f^r(x)$.

Aus diesem Satze folgt, daß eine Iterierte $f^r(x)$ von $f(x)$ *alle* Eigenwerte ζ von $f(x)$ in genau der gleichen Vielfachheit wie $f(x)$ selbst besitzt, wenn für jeden einfachen Eigenwert ζ von $f(x)$ stets $(f'(\zeta))^r \neq 1$ ist. Daher gibt es unendlich viele Iterierte $f^r(x)$, bei denen die Lösungen von $1 + H_r(x) = 0$ keine Eigenwerte von $f(x)$ sind, also $f^r(x)$ neue Eigenwerte hat.

Seien nun $\zeta_1, \zeta_2, \ldots, \zeta_p$ irgendwelche p verschiedenen Eigenwerte von irgendwelchen der f^n, $n \geq 1$, der Folge (4). Sei ζ_λ ein Eigenwert von f^{n_λ}, dann hat auch jede Iterierte von f^{n_λ}, d.h. jedes $f^{r \cdot n_\lambda}$ auch den Eigenwert ζ_λ. Ist nun m das kleinste gemeinschaftliche Vielfache der p Exponenten n_λ, so ist jedes ζ_λ, $\lambda = 1, 2, \ldots, p$, Eigenwert der Iterierten f^m. Da es aber stets für ein Polynom eines Grades > 1 eine Iterierte mit neuen Eigenwerten gibt, hat ein passendes $f^{r \cdot m}$ einen Eigenwert $\zeta \neq \zeta_\lambda$, $\lambda = 1, \ldots, p$. Daher bilden die verschiedenen Eigenwerte der Polynome $f, f^2, f^3, \ldots$ eine abzählbare Folge

$$\zeta_1, \zeta_2, \zeta_3, \ldots, \tag{8}$$

wenn der Grad von $f(x)$ größer als 1 ist. Man schreibt erst alle verschiedenen Eigenwerte von $f(x)$, dann die neuen von $f^2(x)$ usw. auf, wobei es vorkommen kann, daß einige Abteilungen leer sind.

§ 2. Die regulären und die singulären Zahlen eines Polynoms.

Sei $f(x)$ ein Polynom positiven Grades m. Setzt man in der Folge (4) überall für x eine Zahl ϑ ein, so erhält man die Zahlfolge

$$V_{f(\vartheta)} = \{f^n(\vartheta)\}_{n=0,1,2\ldots} = \{\vartheta, f(\vartheta), f^2(\vartheta), \ldots\}. \tag{4a}$$

Sind alle diese Zahlen voneinander verschieden, so soll ϑ eine *reguläre Zahl* des Polynoms $f(x)$ heißen. Jede nicht reguläre Zahl soll eine *singuläre Zahl* des Polynoms genannt werden. Für eine singuläre Zahl ϑ sind also mindestens zwei Elemente der Folge (4a) einander gleich. Für $f(x) \equiv x$ ist jede Zahl singulär.

Für $f(x) = x + b$, $b \neq 0$, ist jede Zahl regulär. Das Polynom $f(x) = ax + b$, a eine von 1 verschiedene Einheitswurzel, hat nur singuläre Zahlen, da ja (4) eine periodische Folge ist. Ist aber a keine Einheitswurzel, so ist $x_0 = \frac{b}{1-a}$ die einzige singuläre Zahl des Polynoms, alle übrigen sind reguläre Zahlen. Sei nun der Grad $m > 1$. Alle Polynome in (4) sind voneinander verschieden. Eine singuläre Zahl ϑ des Polynoms $f(x)$ ist dadurch charakterisiert, daß es eine *kleinste* natürliche Zahl p gibt, für die $f^p(\vartheta)$ gleich einem $f^r(\vartheta)$ mit $r < p$ gibt. Dann sind $f^0(\vartheta) = \vartheta$, $f^1(\vartheta), \ldots, f^{p-1}(\vartheta)$ alle voneinander verschieden. Setzt man $p = r + s$, so folgt aus $f^r(\vartheta) = f^p(\vartheta)$, daß $f^{r+1}(\vartheta) = f^{p+1}(\vartheta)$, $f^{r+2}(\vartheta) = f^{p+2}(\vartheta)$, ... wird. Die s Zahlen

$$f^r(\vartheta), f^{r+1}(\vartheta), \ldots, f^{r+s-1}(\vartheta)$$

wiederholen sich periodisch. Vor ihnen steht die Vorperiode:

$$f^0(\vartheta) = \vartheta, f^1(\vartheta), \ldots, f^{r-1}(\vartheta),$$

die wegfällt, wenn $r = 0$ ist. Die Folge (4a) enthält nur p verschiedene Zahlen. Schreibt man die Gleichung $f^p(\vartheta) = f^r(\vartheta)$, die ϑ als singuläre Zahl des Polynomes $f(x)$ charakterisiert, in der Gestalt

$$f^s(f^r(\vartheta)) = f^r(\vartheta), \tag{9}$$

so sieht man, daß $f^r(\vartheta)$ ein Eigenwert des Polynoms $f^s(x)$ ist, also eine der Zahlen ζ_λ der Folge (8). Man bekommt umgekehrt durch Lösung irgendeiner Gleichung $f^r(x) = \zeta_\lambda$ mit jeder Lösung ϑ eine singuläre Zahl des Polynoms, denn ist ζ_λ ein Eigenwert der s-ten Iterierten, so ist wegen

$$f^r(\vartheta) = \zeta_\lambda, \quad f^s(\zeta_\lambda) = \zeta_\lambda$$
$$f^{s+r}(\vartheta) = f^s(\zeta_\lambda) = \zeta_\lambda = f^r(\vartheta),$$

also ϑ eine singuläre Zahl des Polynoms.

Hieraus folgt

Satz 4. *Die Menge der singulären Zahlen eines Polynoms des Grades $m > 1$ ist abzählbar. Man erhält diese Zahlen ϑ als Lösungen der algebraischen Gleichungen $f^r(\vartheta) = \zeta$, worin ζ jede Zahl der Folge* (8) *bedeutet. Für jedes singuläre ϑ ist die Folge* (4a) *periodisch.*

Nennt man die Anzahl p der verschiedenen Zahlen in (4a) die *Höhe* der singulären Zahl ϑ mit Bezug auf das Polynom $f(x)$, so sind die singulären Zahlen der Höhe 1 gerade die Eigenwerte von $f(x)$ selbst. Die singulären Zahlen der Höhe 2 sind die verschiedenen Lösungen der beiden Gleichungen $f^2(x) = x$, $f^2(x) = f(x)$ unter Weglassung der Eigenwerte von $f(x)$. Dabei kann die Gleichung $f^2(x) = f(x)$ durch $f(x) = \zeta$ ersetzt werden, wobei ζ einen Eigenwert von $f(x)$ bedeutet. Man kann die singulären Zahlen eines Polynoms durch Ordnen nach ihrer Höhe numerieren. Ich bemerke noch, daß für jedes gerade Polynom und für jeden seiner von Null verschiedenen Eigenwerte ζ stets $-\zeta$ eine singuläre Zahl der Höhe 2 ist, da ja dann (4a) die Form $-\zeta, \zeta, \zeta, \zeta, \ldots$ bekommt. Ganz allgemein ist für jeden Eigenwert ζ eines Polynoms $f(x)$ jede von ζ verschiedene Lösung ϑ der Gleichung $f(x) = \zeta$ aus demselben Grunde eine singuläre Zahl der Höhe 2. Ich nenne noch die ziemlich selbstverständlichen Sätze

Satz 5. *Jede reguläre (singuläre) Zahl des Polynoms $f(x)$ ist auch eine reguläre (singuläre) Zahl jeder Iterierten von $f(x)$ und umgekehrt.*

Satz 6. *Ist ϑ regulär (singulär) mit Bezug auf das Polynom $f(x)$, so ist auch jedes $f^r(\vartheta)$ regulär (singulär) mit Bezug auf $f(x)$.*

Jedes Polynom führt also zu einer Einteilung aller Zahlen in 2 Klassen S_f und R_f, die der singulären und die der regulären Zahlen, und jede Iterierte f^r, $r \geq 1$, liefert dieselbe Einteilung (Satz 5). Die Frage nach allen Polynomen, die dieselbe Einteilung wie ein gegebenes Polynom $f(x)$ liefern, wird später gestreift werden.

§ 3. Vertauschbare Polynome.

Bekannt ist

Satz 7. *Ist ein Polynom $f(x)$ sowohl mit $g(x)$ als auch mit $h(x)$ vertauschbar, so ist f auch mit gh und mit hg vertauschbar.*

Wenn g und h beide mit f vertauschbar sind, brauchen aber g und h nicht miteinander vertauschbar zu sein. Es gibt aber einige Klassen von Polynomen f, für die für je zwei mit f vertauschbare Polynome g und h auch $gh = hg$ folgt.

Definition 1. Eine Menge von Polynomen heiße eine V-Menge, wenn alle ihre Elemente miteinander vertauschbar sind. Ist die Menge speziell abzählbar und erscheint sie in der numerierten Form $f_1, f_2, \ldots$, so heiße sie eine V-Folge.

Die Folge V_f in (4) ist abgesehen von den früher erwähnten Fällen, in denen V_f periodisch war, eine V-Folge, also sicher, wenn der Grad von $f > 1$ ist.

Definition 2. Eine V-Menge (Folge) heiße *abgerundet*, wenn sie mit g und h auch gh enthält.

Ist also der Grad von f größer als 1, so ist V_f eine abgerundete V-Folge. Enthält eine abgerundete V-Folge das Polynom f, so enthält sie auch alle Elemente von V_f, abgesehen eventuell von $f^0 = x$.

Definition 3. Eine V-Menge (Folge) heiße vollständig, wenn es außerhalb der Menge (Folge) kein Polynom *positiven* Grades gibt, das mit jedem Element der Menge (Folge) vertauschbar ist.

Eine V-Menge (Folge) enthält höchstens ein Polynom 0-ten Grades, da zwei voneinander verschiedene Polynome 0-ten Grades $f(x) \equiv a$, $g(x) \equiv b \neq a$ nie miteinander vertauschbar sind. Enthält eine V-Menge (Folge) ein Polynom 0-ten Grades $g(x) \equiv a$, so ist a nach Satz 2 Eigenwert jedes Polynoms der Menge (Folge).

Satz 8. *Eine vollständige V-Menge (Folge) ist stets abgerundet.*

Beweis. Liegen g und h in der Menge, so ist gh mit jedem Element der Menge vertauschbar (Satz 7), also liegt wegen der Vollständigkeit der Menge gh in der Menge, wenn der Grad von gh größer als 0 ist. Ist aber der Grad von gh gleich 0, also gh eine Konstante a, so ist g oder h oder beide gleich der einzigen in der Menge vorkommenden Konstanten a, also liegt gh stets in der Menge, die also abgerundet ist.

Definition 4. Eine V-Folge von Polynomen positiven Grades heiße eine V-Kette, wenn in ihr Polynome *aller* positiven Grade vorkommen.

Zum Beispiel ist die Folge

$$x, x^2, x^3, \ldots$$

eine V-Kette.

Wir werden später sehen, daß eine V-Kette nie zwei Polynome gleichen Grades enthält, also jeder Grad nur einmal vertreten ist. Daraus folgt dann sofort die Vollständigkeit der Kette. Wir werden ferner zeigen, wie man alle Ketten finden kann.

Hat die Gl. (1) bei gegebenem f nur die Lösungen $g = f^r$, so ist V_f eine vollständige V-Folge, und umgekehrt folgt aus der Vollständigkeit von V_f, daß die einzigen mit f vertauschbaren Polynome die Iterierten von f sind.

Satz 9. *Die Menge M aller Polynome $x + a$ ist eine vollständige V-Menge. Sie hat die Mächtigkeit des Kontinuums.*

Beweis. Für zwei Polynome $f(x) = x + a$, $g(x) = x + b$ der Menge ist $fg = gf = x + a + b$. Also ist M eine V-Menge, die abgerundet ist. Daß M vollständig ist, folgt schon allein daraus, daß ein mit $g(x) = x + 1$ vertauschbares Polynom $f(x)$ zur Menge gehört. Denn wenn $f(x)$ mit $x + 1$ vertauschbar ist, so ist $f(x)$ von positivem Grad, da $x + 1$ mit keiner Konstanten vertauschbar ist, und es ist $f(x + 1) = f(x) + 1$, also hat $f(x)$ den Grad 1,

$$f(x) = c\,x + d$$

$$f(x + 1) = c(x + 1) + d = f(x) + 1 = c\,x + d + 1$$

$$c = 1$$

$$f(x) = x + d.$$

Also gehört $f(x)$ zu M. M ist vollständig.

Satz 10 (Umkehrung von Satz 1). *Das einzige Polynom, das mit jedem Polynom vertauschbar ist, ist $g(x) = x$.*

Beweis. Ist $g(x)$ mit jedem Polynom vertauschbar, so ist es unter anderem mit $f(x) = x + 1$ und $h(x) = 1$ vertauschbar. Aus $gf = fg$ folgt nach dem Beweise von Satz 9, daß $g(x) = x + a$ und aus $gh = hg$ oder $1 + a = 1$, $a = 0$, also $g(x) = x$.

Satz 11. *Sind $f(x)$ und $g(x)$ miteinander vertauschbar und ist ζ ein Eigenwert von $g(x)$, so ist auch $f(\zeta)$ ein Eigenwert von $g(x)$. Ist ferner η eine Lösung der Gleichung $f(x) = \zeta$, so ist auch $g(\eta)$ eine Lösung dieser Gleichung.*

Beweis. Aus $g(\zeta) = \zeta$ folgt $f(\zeta) = fg(\zeta) = g(f(\zeta))$. Daher ist $f(\zeta)$ auch ein Eigenwert von $g(x)$. — Aus $f(\eta) = \zeta$ folgt $\zeta = g(\zeta) = gf(\eta) = f(g(\eta))$. Also löst $g(\eta)$ die Gleichung $f(x) = \zeta$.

Der letzte Teil werde am Beispiel $g(x) = x^2 - 2$, $f(x) = x^3 - 3\,x$ illustriert, die vertauschbare Polynome sind. Nehmen wir als Eigenwert von $g(x)$ etwa $\zeta = -1$, so ist die Gleichung $f(\eta) = \zeta$: $\eta^3 - 3\eta + 1 = 0$. Dann ist nach dem letzten Teil des Satzes mit jeder Lösung η auch $\eta^2 - 2$ eine Lösung. Die kubische Gleichung ist eine Normalgleichung, ohne daß eine Adjunktion nötig ist.

Sie hat drei reelle Lösungen: $\eta_1 = 0{,}35\ldots$; $\eta_2 = -0{,}87\ldots$; $\eta_3 = 1{,}53\ldots$. $\eta_2 = \eta_1^2 - 2 = g(\eta_1)$, $\eta_3 = g(\eta_2) = g^2(\eta_1)$, $\eta_1 = g(\eta_3) = g^3(\eta_1)$. Also $\eta_\lambda = g^3(\eta_\lambda)$. Es wird mit

$$\varphi(x) = f(x) + 1 = x^3 - 3\,x + 1:$$

$$g^3(x) - x = \varphi(x)\,(x^5 - 5\,x^3 - x^2 + 5\,x + 2)$$

$$\varphi(x^2 - 2) = \varphi(x)\,(\varphi(x) + 2).$$

Satz 5 lehrte, daß die zu einem Polynom gehörende Einteilung aller Zahlen in 2 Klassen R und S von regulären und singulären Zahlen für jede Iterierte des Polynomes die gleiche Einteilung ergibt. Das ist ein spezieller Fall von

Satz 12. *Zwei vertauschbare Polynome $f(x)$ und $g(x)$, deren Grade größer als 1 sind, bestimmen dieselbe Einteilung aller Zahlen in zwei Klassen R und S, d. h. ist ϑ eine singuläre Zahl für f, so ist ϑ auch eine singuläre Zahl des Polynomes g und umgekehrt. Daraus folgt dann das gleiche für die regulären Zahlen.*

Beweis. Wegen der Symmetrie der Voraussetzung in f und g genügt es zu zeigen: Jede singuläre Zahl von g ist auch singuläre Zahl des Polynomes f. Da $g(x)$ einen Grad >1 hat, besitzt $g(x)$ singuläre Zahlen. ϑ sei irgendeine von ihnen. Dann ist für passendes r, s, $s \geqq 1$,

$$g^r(\vartheta) = g^{r+s}(\vartheta) = g^s g^r(\vartheta),$$

also $g^r(\vartheta)$ Eigenwert von $g^s(x)$. Da f mit g vertauschbar ist, so ist auch $f^n(x)$ mit $g^s(x)$ vertauschbar, also nach Satz 11 $f^n g^r(\vartheta) = g^r f^n(\vartheta)$, auch ein Eigenwert von $g^s(x)$. Sind alle verschiedenen Eigenwerte von $g^s(x)$ etwa $x_1, \ldots, x_q$ und sind alle verschiedenen Lösungen der q Gleichungen

$$g^r(t) = x_\lambda, \qquad \lambda = 1, \ldots, q,$$

etwa $t_1, \ldots, t_m$, so ist also wegen $g^r\left(f^n(\vartheta)\right) = x_\lambda$ die Zahl $f^n(\vartheta)$ gleich einem der t_μ für jedes $n = 1, 2, 3, \ldots$. Daher besteht $V_{f(\vartheta)}$ nur aus endlich vielen verschiedenen Zahlen. ϑ ist also singuläre Zahl des Polynoms f, w. z. b. w.

Für den Fall, daß eines der Polynome oder beide linear sind, braucht der Satz nicht zu gelten. Das läßt sich aber besser übersehen, wenn man erst alle mit einem gegebenen linearen Polynom $g(x) = ax + b$ vertauschbaren Polynome bestimmt hat.

§ 4. Das lineare Polynom $g(x) = ax + b$.

Will man alle mit $g(x) = ax + b$ vertauschbaren Polynome $f(x)$ bestimmen, kann man $a = 1$ beiseite lassen (Satz 9). Die Bedingung dafür, daß ein Polynom $f(x)$ mit $g(x) = ax + b$ vertauschbar ist, lautet

$$f(ax + b) = af(x) + b, \tag{1'}$$

wobei wir also $a \neq 1$ annehmen können.

Der Eigenwert von $g(x)$ ist

$$\zeta = \frac{b}{1 - a}, \tag{10}$$

also

$$g(x) = a(x - \zeta) + \zeta, \quad a\zeta + b = \zeta. \tag{10a}$$

Aus (1') oder Satz 11 folgt für jedes mit $g(x)$ vertauschbare Polynom $f(x)$, daß $f(\zeta) = \zeta$, also $f(x)$ auch ζ zum Eigenwert hat. Setzt man nun

$$x = y + \zeta, \quad y = x - \zeta, \tag{11}$$

so wird

(11a) $$f(x)-\zeta=f(y+\zeta)-\zeta=h(y)$$

ein Polynom in y, das $y=0$ zum Eigenwert hat.

Daher ist

(12) $$h(y)=y\,k(y),$$

worin $k(y)$ den Grad $(n-1)$ hat, wenn $f(x)$ vom Grade n ist. Indem man nun in (1') mittels (11), (11a) die neue veränderliche y einführt und (10) und (12) benützt, erhält man die mit (1') gleichwertige Gleichung

(13) $$k(a\,y)=k(y).$$

Der Fall, in dem $f(x)$ eine Konstante ist, also $f(x)\equiv\zeta'$, $h(y)\equiv 0$ ist, ist als trivial beiseite gelassen worden. Ist $f(x)$ linear, so ist $k(y)$ eine Konstante C,

$$h(y)=C\,y=C(x-\zeta)=f(x)-\zeta$$

(14) $$f(x)=C(x-\zeta)+\zeta,$$

was für $C=0$ den eben erwähnten Fall $f(x)\equiv\zeta$ ergibt. ζ ist auch Eigenwert jedes linearen mit $g(x)$ vertauschbaren Polynomes. Die durch (14) definierten Polynome, unter denen für $C=a$ auch $g(x)$ vorkommt, sind alle untereinander vertauschbar.

Es gilt nämlich

Satz 13. *Die Menge aller Polynome*

$$f(x)=C(x-\zeta)+\zeta$$

bildet bei gegebenem festen ζ eine abgerundete V-Menge.

Die Richtigkeit folgt unmittelbar aus den Gleichungen

$$f_1(x)=C_1(x-\zeta)+\zeta,$$
$$f_2(x)=C_2(x-\zeta)+\zeta,$$
$$f_1f_2=C_1C_2(x-\zeta)+\zeta=f_2f_1.$$

Die durch ζ bestimmte Polynommenge M_ζ hat Kontinuummächtigkeit. Jedes Element in M_ζ hat ζ als Eigenwert, der für $C=0$ selbst zur Menge gehört.

Sei nun der Grad von $f(x)$ größer als 1, dann lehrt der Vergleich der Koeffizienten von y^{n-1} in (13), daß

(15) $$a^{n-1}=1$$

ist. Ist also a keine Einheitswurzel, so ist der Fall $n>1$ unmöglich. Hieraus folgt

Satz 14. *Die Menge M_ζ des Satzes 13 ist sogar vollständig.*

Beweis. Es gibt kein Polynom eines Grades $n\geqq 2$, das mit jedem Polynom in M_ζ vertauschbar ist, denn unter den Polynomen in M_ζ kommt z.B. auch das Polynom

$$2(z-\zeta)+\zeta=2z-\zeta$$

vor, und da 2 keine Einheitswurzel ist, so gibt es kein Polynom eines Grades $n \geq 2$, das mit $2z-\zeta$ vertauschbar ist.

Ist aber ein Polynom $h(x)$ linear und mit auch nur einem Element $C(x-\zeta)+\zeta$, $C \neq 1$, in M_ζ vertauschbar, so ist nach Satz 11 auch $h(\zeta)$ ein Eigenwert von $C(x-\zeta)+\zeta$, d.h. $h(\zeta)=\zeta$ und dann hat $h(x)$ die Gestalt $C'(x-\zeta)+\zeta$ und liegt in M_ζ. Außerhalb M_ζ gibt es also kein Polynom positiven Grades, das mit allen Polynomen in M_ζ vertauschbar ist. Daher ist M_ζ eine vollständige V-Menge. Mit Rücksicht auf die im Anschluß an Satz 7 gemachten Bemerkungen sei hervorgehoben: Sind $f(x)$ und $h(x)$ beide mit $g(x)=ax+b$ vertauschbar, worin a keine Einheitswurzel ist, dann ist auch $f(x)$ mit $h(x)$ vertauschbar. Dies gilt auch noch für $a=1$, $b \neq 0$ (s. Satz 9).

Nun sei $a \neq 1$ eine primitive r-te Einheitswurzel, dann ist nach (15)

$$n \equiv 1 \,(r), \qquad n = 1 + \lambda r. \tag{16}$$

Wegen

$$k(y) = \sum_{\nu=0}^{n-1} b_\nu y^\nu \tag{17}$$

folgt aus (13), daß in (17) nur solche Glieder vorkommen, bei denen $\nu \equiv 0 \,(r)$ ist; es ist also $k(y)$ ein Polynom von y^r, also mit neuer Bezeichnung

$$k(y) = \sum_{k=0}^{\lambda} a_k y^{rk}. \tag{18}$$

Daher wird nach (11a) und (12)

$$f(x) = (x-\zeta) \sum_{k=0}^{\lambda} a_k (x-\zeta)^{rk} + \zeta. \tag{19}$$

Daß diese Polynome wirklich mit $g(x)=ax+b$ vertauschbar sind, rechnet man unter Benutzung von $a^r=1$, $(r \geq 2)$, leicht nach. Wegen der Willkürlichkeit der Koeffizienten a_k bilden die Polynome (19) eine Menge von der Mächtigkeit des Kontinuums. Wir haben also

Satz 15. *Ist $g(x)=ax+b$, worin a eine primitive r-te Einheitswurzel $(r \geq 2)$ ist, und ist ζ der Eigenwert von $g(x)$, so ist die Gesamtheit aller mit $g(x)$ vertauschbaren Polynome $f(x)$ durch* (19) *gegeben. Für $\lambda=0$ erhält man speziell die früher betrachteten linearen Polynome der Menge M_ζ des Satzes 13.*

§ 5. Ähnliche Polynome.

Ist ein Polynom $f(x)$ mit dem linearen Polynom $ax+b$, $a \neq 0$, vertauschbar, also

$$f(ax+b) = af(x) + b,$$

so ist

$$\frac{f(ax+b)-b}{a} = f(x).$$

Ist $f(x)$ mit $ax+b$ nicht vertauschbar, so stellt

$$\frac{f(ax+b)-b}{a} = f_1(x) \tag{20}$$

ein von $f(x)$ verschiedenes Polynom dar, das aus $f(x)$ durch eine Ähnlichkeitstransformation gewonnen wird. Die Matrix

$$M = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \tag{21}$$

definiert die lineare Transformation:

$$x = a\,x' + b.$$

Transformiert man für $y = f(x)$ kogredient

$$y = a\,y' + b,$$

so bekommt man:

$$y' = \frac{f(a\,x' + b) - b}{a} = f_1(x'). \tag{20a}$$

Bezeichnet man die ganze lineare Transformation mit

$$T = T(a, b), \tag{22}$$

so ist

$$T^{-1} = T\left(\frac{1}{a}, -\frac{b}{a}\right) \tag{22a}$$

und der Zusammenhang zwischen $f(x)$ und $f_1(x)$ schreibt sich einfach

$$f_1(x) = T^{-1} f\, T(x). \tag{23}$$

$f_1(x)$ heißt ein zu $f(x)$ ähnliches Polynom oder die T-Transformierte von $f(x)$. Dann gilt

Satz 16. *Sind $f(x)$ und $g(x)$ zwei Polynome und f_1 und g_1 ihre T-Transformierten, dann ist $f_1 g_1$ die T-Transformierte von fg.*

Denn $f_1 g_1 = T^{-1} f T g_1(x) = T^{-1} f T\, T^{-1} g\, T = T^{-1}(fg)\, T$. Daraus folgt sofort: Ist $fg = gf$, so ist auch $f_1 g_1 = g_1 f_1$ oder

Satz 17. *Transformiert man zwei vertauschbare Polynome $f(x)$ und $g(x)$ durch dieselbe Transformation T in die zu ihnen ähnlichen Polynome $f_1(x)$, $g_1(x)$, so ist auch f_1 und g_1 vertauschbar.*

Will man also alle Polynome $f(x)$ bestimmen, die mit einem gegebenen Polynom $g(x)$ eines Grades ≥ 2 vertauschbar sind, so ist es keine Einschränkung der Allgemeinheit, wenn man $g(x)$ normiert annimmt und den zweithöchsten Koeffizienten von $g(x)$ gleich 0 setzt. Dann wendet man auf $g(x)$ und das gesuchte $f(x)$ eine und dieselbe T-Transformation $T(a, b)$ an, so ist

$$g_1 = \frac{g(ax + b) - b}{a} \quad \text{auch mit} \quad f_1 = \frac{f(a\,x + b) - b}{a}$$

vertauschbar, und wenn

$$g(x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_n$$

ist, so wird

$$g_1 = \frac{b_0(a\,x+b)^n + b_1(a\,x+b)^{n-1} + \cdots + b_n - b}{a}$$

$$= \frac{b_0 a^n x^n + (b_0 n a^{n-1} b + b_1 a^{n-1})\, x^{n-1} + \cdots + b_n - b}{a}$$

$$= b_0 a^{n-1} x^n + a^{n-2}(n\,b\,b_0 + b_1)\, x^{n-1} + \cdots.$$

Da $n \geqq 2$ ist, so geben die Gleichungen

$$b_0 a^{n-1} = 1, \qquad a^{n-2}(n\,b\,b_0 + b_1) = 0$$

für a und b die Werte

$$a = \sqrt[n-1]{\frac{1}{b_0}}, \qquad b = -\frac{1}{n}\frac{b_1}{b_0}$$

für die gewünschte Transformation. Ist $n = 2$, so ist die Bestimmung von a und b *eindeutig*.

Ist ein Polynom $g(x)$ vom Grade n gegeben, und sind seine Eigenwerte $\zeta_1, \zeta_2, \ldots, \zeta_n$, so hat das durch $T(a, b)$ erhaltene ähnliche Polynom

$$g_1(x) = \frac{g(a\,x+b) - b}{a}$$

die Eigenwerte $\frac{\zeta_\lambda - b}{a}$, $\lambda = 1, 2, \ldots, n$. Denn die Gleichung für die Eigenwerte von $g_1(x)$ lautet $g_1(x) = x$ oder

$$g(a\,x+b) = a\,x + b,$$

die die Lösungen $a\,x + b = \zeta_\lambda$, $x = \frac{\zeta_\lambda - b}{a} = \vartheta_\lambda$ ergibt. Ist speziell $n = 2$, und transformiert man das quadratische Polynom $g(x)$ in ein ähnliches mit demselben höchsten Koeffizienten, d.h. wählt man $a = 1$, so ist

$$\vartheta_1 = \zeta_1 - b$$
$$\vartheta_2 = \zeta_2 - b$$

und $\vartheta_1 - \vartheta_2 = \zeta_1 - \zeta_2$. Die Differenz der Eigenwerte ist also invariant gegenüber einer solchen Transformation.

§ 6. Die Polynome

$$g(x) = b_0 x^n + b_1 x^{n-2} + \ldots + b_n, \qquad n \geqq 2 \qquad b_0 \neq 0.$$

Über die Menge M der Polynome $f(x)$, die mit dem gegebenen Polynom $g(x)$ der Überschrift vertauschbar sind, gilt

Satz 18. *Ist $g(x)$ ein Polynom eines Grades $n \geqq 2$, so bilden alle mit ihm vertauschbaren Polynome $f(x)$ eine abzählbare Menge.*

Dieser Satz ist eine unmittelbare Folge von

Satz 19. *Ist m eine gegebene Gradzahl ≥ 0, so gibt es für $m=0$ höchstens n Polynome $f(x)$ des Grades $m=0$, die mit $g(x)$ vertauschbar sind. Für $m \geq 1$ gibt es höchstens $(n-1)$ Polynome $f(x)$ des Grades m, die mit $g(x)$ vertauschbar sind.*

Beweis. Ist der Grad eines mit $g(x)$ vertauschbaren Polynomes $f(x)$ gleich Null, so ist $f(x)$ eine Konstante, die nach Satz 2 ein Eigenwert von $g(x)$ ist. Und da es höchstens n verschiedene solche gibt, ist der Satz im Falle $m=0$ richtig. — Sei nun $m \geq 1$. Wir betrachten vorweg den Fall $m=1$, in dem $f(x)=a_0 x+a_1$, $a_0 \neq 0$, sein soll. Die Vertauschbarkeitsbedingung lautet: $fg=gf$. Nun ist:

$$fg = a_0 g(x) + a_1 = a_0 \sum_{\nu=0}^{n} b_\nu x^{n-\nu} + a_1$$

$$gf = g(a_0 x + a_1) = \sum_{\nu=0}^{n} b_\nu (a_0 x + a_1)^{n-\nu}.$$

Der Vergleich der Koeffizienten von x^n und x^{n-1} in beiden Ausdrücken gibt

$$b_0 a_0^n = a_0 b_0, \quad a_0^{n-1} = 1$$

$$b_0 n a_0^{n-1} a_1 + b_1 a_0^{n-1} = a_0 b_1$$

$$b_0 n a_1 = b_1 (a_0 - 1)$$

$$a_1 = \frac{b_1 (a_0 - 1)}{n b_0}.$$

a_0 ist $(n-1)$-deutig bestimmt, und zu jedem dieser $(n-1)$ Werte a_0 ist a_1 eindeutig bestimmt. Ob mit den so gefundenen Werten von a_0, a_1 die niederen Koeffizienten der Vertauschbarkeitsgleichung links und rechts übereinstimmen, ist eine offene Frage, jedenfalls kann es nicht mehr als $(n-1)$ Polynome ersten Grades geben, die mit $g(x)$ vertauschbar sind. Da der Eigenwert des Polynoms $f(x)=a_0 x+a_1$, durch

$$\xi = \frac{a_1}{1-a_0} = -\frac{b_1}{n b_0}$$

gegeben ist, ist das gefundene $f(x)$ nur unter gewissen Bedingungen zwischen den b_ν mit $g(x)$ vertauschbar, die aus § 4 zu entnehmen sind. — Sei nun der Grad m von $f(x)$ größer als 1. Es sei

$$f(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_m; \qquad m \geq 2;\ a_0 \neq 0. \tag{24}$$

Die Frage ist die Bestimmung der $(m+1)$ Koeffizienten $a_0, a_1, \ldots, a_m$ mit Hilfe der Gl. (1). Da der Grad in (1) auf beiden Seiten gleich mn ist, sind im ganzen $(mn+1)$ Gleichungen von den $(m+1)$ Unbekannten a_ν zu erfüllen, also mehr Gleichungen als Unbekannte zur Verfügung stehen; es wird also im allgemeinen gar keine Lösung geben. Wir müssen zeigen, daß es, wenn die Gleichungen lösbar sind, höchstens $(n-1)$ Lösungen bei gegebenem m gibt. Die Gl. (1) lautet hier

$$\sum_{\lambda=0}^{m} a_\lambda \big(g(x)\big)^{m-\lambda} = \sum_{\lambda=0}^{n} b_\lambda \big(f(x)\big)^{n-\lambda}. \tag{1*}$$

Um die Koeffizientenvergleichung durchführen zu können, muß man alle Potenzen $(f(x))^p$, $(g(x))^p$ nach Potenzen von x entwickeln. Der polynomische Satz ergibt:

$$(25)\quad \begin{cases} (f(x))^p = \displaystyle\sum_{\lambda_0+\lambda_1+\cdots+\lambda_m=p} \frac{p!}{\lambda_0!\,\lambda_1!\ldots\lambda_m!}\, a_0^{\lambda_0} x^{m\lambda_0} a_2^{\lambda_1} x^{(m-1)\lambda_1} \ldots a_m^{\lambda_m} x^{(m-m)\lambda_m} \\ \quad = \displaystyle\sum_{\lambda_0+\lambda_1+\cdots+\lambda_m=p} \frac{p!}{\lambda_0!\,\lambda_1!\ldots\lambda_m!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_m^{\lambda_m} x^{mp-1\lambda_1-\cdots-m\lambda_m}. \end{cases}$$

Daher wird der Koeffizient $A_{p,k}$ von x^k in $(f(x))^p$

$$(26)\quad A_{p,k} = \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_m=p \\ 1\cdot\lambda_1+2\lambda_2+\cdots+m\lambda_m=mp-k}} \frac{p!}{\lambda_0!\,\lambda_1!\ldots\lambda_m!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_m^{\lambda_m}, \qquad k = 0, 1, 2, \ldots, mp.$$

Entsprechend wird der Koeffizient von x^k in $(g(x))^p$

$$(27)\quad B_{p,k} = \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_n=p \\ 1\cdot\lambda_1+2\lambda_2+\cdots+n\lambda_n=np-k}} \frac{p!}{\lambda_0!\,\lambda_1!\ldots\lambda_n!}\, b_0^{\lambda_0} b_1^{\lambda_1} \ldots b_n^{\lambda_n}, \qquad k = 0, 1, 2, \ldots, np.$$

Der Vergleich der Koeffizienten von x^{mn} in (1*) ergibt

$$(28)\quad a_0 b_0^m = b_0 a_0^n$$

oder

$$(28_1)\quad a_0^{n-1} = b_0^{m-1}.$$

Hierdurch ist der höchste Koeffizient a_0 von $f(x)$ $(n-1)$-deutig bestimmt. Wenn die weiteren Koeffizienten $a_1, a_2, \ldots, a_m$ sich rekursiv durch lineare Gleichungen bestimmen lassen, falls man einen der gefundenen $(n-1)$ Werte a_0 zugrunde legt, dann ist damit gezeigt, daß es höchstens $(n-1)$ Polynome $f(x)$ des Grades m geben kann, die (1*) befriedigen. Nun vergleichen wir in (1*) die Koeffizienten von x^{mn-1} auf beiden Seiten. Da m und n beide >1 sind, kommt diese Potenz links nur in $a_0(g(x))^m$ und rechts nur in $b_0(f(x))^n$ vor. Da die Koeffizienten von x^{mn-1} aus (27) und (26) zu entnehmen sind, gibt der Vergleich

$$a_0 B_{m,mn-1} = b_0 A_{n,mn-1}$$

oder

$$a_0 \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_n=m \\ 1\cdot\lambda_1+2\lambda_2+\cdots+n\lambda_n=1}} \frac{m!}{\lambda_0!\,\lambda_1!\ldots\lambda_n!}\, b_0^{\lambda_0} b_1^{\lambda_1} \ldots b_n^{\lambda_n} = b_0 \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_m=n \\ 1\cdot\lambda_1+2\lambda_2+\cdots+m\lambda_m=1}} \frac{n!}{\lambda_0!\,\lambda_1!\ldots\lambda_m!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_m^{\lambda_m}.$$

Dies reduziert sich wegen der unter den Summen stehenden Relationen zwischen den Summationsbuchstaben auf

$$a_0 m b_0^{m-1} b_1 = b_0 n a_0^{n-1} a_1 = b_0 n b_0^{m-1} a_1,$$

wenn man (28_1) benutzt. Daraus folgt

$$(29)\quad a_1 = \frac{m}{n}\,\frac{a_0 b_1}{b_0},$$

so daß zu jedem der $(n-1)$ erhaltenen Werte a_0 der nächste Koeffizient a_1 eindeutig bestimmt ist.

Nehmen wir nun an, wir hätten schon mit Benutzung der gefundenen Werte a_0, a_1 durch Vergleich der Koeffizienten von x^{mn-2} den nächsten Koeffizienten a_2 eindeutig gefunden und weiter durch Vergleich der Koeffizienten von

$$x^{mn-3}, x^{mn-4}, \ldots, x^{mn-q}$$

in (1*) die Koeffizienten

$$a_3, a_4, \ldots, a_q, \qquad q \leq m-1,$$

eindeutig errechnet, so daß wir also $a_0, a_1, \ldots, a_q$ als schon bekannt ansehen können. Wenn sich dann auch a_{q+1} durch den Vergleich der Koeffizienten von

$$x^{mn-q-1}$$

in (1*) eindeutig ergibt, dann folgt durch Induktion, daß zu jedem der $(n-1)$ Werte a_0, die sich aus (28_1) ergeben, $a_1, a_2, \ldots, a_m$ eindeutig bestimmt sind, es also höchstens $(n-1)$ Polynome $f(x)$ des Grades m geben kann, die mit $g(x)$ vertauschbar sind. Wir haben also zu zeigen, daß der Vergleich der Koeffizienten von x^{mn-q-1} in (1*) wirklich auf eine Gleichung ersten Grades für a_{q+1} führt.

Wegen $q \leq m-1$ ist $-(q+1) \geq -m$ und

$$mn-(q+1) \geq mn-m = m(n-1).$$

Gilt hier das Zeichen $>$, was für $q < m-1$ der Fall ist, dann kommen Glieder mit dem Exponenten $mn-(q+1)$ auf der rechten Seite in (1*) nur in $b_0(f(x))^n$ vor, da ja das nächste Glied nur noch einen Grad $m(n-1) < mn-(q+1)$ hat. Nur für $q = m-1$ liefert dieses nächste Glied noch einen Beitrag mit dem Exponenten $mn-(q+1) = m(n-1)$, nämlich das Glied

$$b_1 a_0^{n-1} x^{m(n-1)},$$

das aber a_{q+1} nicht enthält. Das gesuchte a_{q+1} kommt also auf der rechten Seite von (1*) nur in den Gliedern vor, die von $b_0(f(x))^n$ stammen. In diesem Ausdruck wird der Koeffizient von $x^{mn-(q+1)}$ gleich

$$b_0 A_{n,mn-(q+1)} = b_0 \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_m=n \\ 1\cdot\lambda_1+2\lambda_2+\cdots+m\lambda_m=q+1}} \frac{n!}{\lambda_0!\,\lambda_1!\ldots\lambda_m!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_m^{\lambda_m}.$$

Da nun wegen der Relationen zwischen den λ_ϱ

$$\varrho\,\lambda_\varrho \leq q+1, \qquad \lambda_\varrho \leq \frac{q+1}{\varrho}$$

ist, so ist für $\varrho > q+1$ sicher $\lambda_\varrho = 0$. Daher reduziert sich die letzte Formel auf

$$b_0 A_{n,mn-(q+1)} = b_0 \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_{q+1}=n \\ 1\cdot\lambda_1+2\lambda_2+\cdots+(q+1)\lambda_{q+1}=q+1}} \frac{n!}{\lambda_0!\,\lambda_1!\ldots\lambda_{q+1}!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_{q+1}^{\lambda_{q+1}}.$$

Es tritt also nur ein Glied in der Summe auf, das a_{q+1} enthält, und zwar in der ersten Potenz, nämlich für

$$\lambda_{q+1} = 1, \quad \lambda_1 = \lambda_2 = \cdots = \lambda_q = 0, \quad \lambda_0 = n - 1.$$

Alle anderen noch auftretenden Glieder der Summe entstehen nur für $\lambda_{q+1} = 0$, d.h. sie enthalten a_{q+1} nicht.

Der Koeffizient von x^{mn-q-1} rechts in (1*) wird also

$$b_0 \left[n\, a_0^{n-1} a_{q+1} + \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_q=n \\ 1\cdot\lambda_1+\lambda_2+\cdots+q\,\lambda_q=q+1}} \frac{n!}{\lambda_0!\,\lambda_1!\ldots\lambda_q!}\, a_0^{\lambda_0} a_1^{\lambda_1} \ldots a_q^{\lambda_q} \right],$$

wozu noch im Falle $q = m - 1$ das Glied

$$b_1 a_0^{n-1}$$

zuzufügen ist.

Auf der linken Seite von (1*) hat das Glied

$$a_\varrho \big(g(x)\big)^{m-\varrho} = a_\varrho (b_0 x^n + \cdots + b_n)^{m-\varrho}$$

den Grad $n(m-\varrho)$ und dieser ist $\geq mn - (q+1)$ nur für

$$\varrho n \leq q + 1;$$

wegen $n \geq 2$ folgt, daß dies nur der Fall ist für

$$\varrho \leq \frac{q+1}{2} \leq q,$$

da ja $q \geq 1$ ist. Der Koeffizient von x^{mn-q-1} auf der linken Seite von (1*), der ja in den a_λ linear ist, enthält also höchstens die schon bekannten Koeffizienten

$$a_0, a_1, \ldots, a_q.$$

Daher bekommen wir wirklich durch Vergleich der Koeffizienten von $x^{mn-(q+1)}$ in (1*) eine in a_{q+1} lineare Gleichung, die a_{q+1} eindeutig bestimmt, so daß der Satz bewiesen ist[2]).

Speziell für $n = 2$ folgt aus diesem Satze

Satz 20. *Ist $g(x)$ ein Polynom des Grades 2, so gibt es höchstens ein Polynom $f(x)$ eines gegebenen positiven Grades m, das mit $g(x)$ vertauschbar ist.*

Und hieraus folgt weiter

Satz 21. *Sind $f_1(x)$ und $f_2(x)$ zwei Polynome positiven Grades, die beide mit einem Polynom $g(x)$ des Grades 2 vertauschbar sind, so ist $f_1(x)$ auch mit $f_2(x)$ vertauschbar.*

[2]) Man kann mit Hilfe des zum Beweise von Satz 19 benutzten Induktionsverfahrens zeigen: Ist $g(x)$ ein gerades oder ungerades Polynom eines Grades $n \geq 2$, so ist jedes mit $g(x)$ vertauschbare Polynom $f(x)$ entweder ein gerades oder ein ungerades Polynom: Für den Fall, daß $g(x)$ ein gerades Polynom ist, läßt sich die Gleichung $f(-x) = \pm f(x)$ leicht direkt aus $fg = gf$ ableiten. Bei ungeradem n dagegen versagt diese Schlußweise, weil der Satz im Falle $n = 1$, $g(x) = ax$, falsch ist. Mit $g(x) = x$ ist jedes Polynom vertauschbar; mit $g(x) = \omega x$, $\omega^3 = 1$, $\omega \neq 1$, ist $f(x) = x + x^4$ vertauschbar.

Beweis. Da $f_1 f_2$ und $f_2 f_1$ beide nach Satz 7 in § 3 ebenfalls mit $g(x)$ vertauschbar sind und $f_1 f_2$ und $f_2 f_1$ gleichen positiven Grad haben, so ist nach Satz 20: $f_1 f_2 = f_2 f_1$.

Als Illustration zu Satz 19 sei erwähnt:

Ist $g(x) = x^n$, $n \geq 2$, so gibt es genau $(n-1)$ Polynome eines gegebenen positiven Grades m, die mit x^n vertauschbar sind. Es sind das nämlich die $(n-1)$ Polynome $C x^m$, worin C jede der Einheitswurzeln bedeutet, die durch $C^{n-1} = 1$ gegeben sind.

Aus Satz 20 folgt noch für die in § 3, Definition 4, erklärten V-Ketten

Satz 22. *Jede V-Kette enthält genau ein Polynom m-ten Grades, $m = 1, 2, 3 \ldots$ und ist stets vollständig.*

Beweis. Da eine solche Kette ein Polynom $g(x)$ des Grades 2 enthält und es zu jedem Grade $m \geq 1$ höchstens ein Polynom $f(x)$ des Grades m gibt, das mit $g(x)$ vertauschbar ist, die Kette aber laut Definition Polynome aller positiven Grade enthält, so enthält die Kette für jeden positiven Grad nur ein Polynom dieses Grades, und außerhalb der Kette existiert kein anderes Polynom dieses Grades, das mit $g(x)$ vertauschbar ist. Also ist die Kette vollständig. So ist die früher in § 3 genannte Folge

$$x, x^2, x^3, \ldots$$

eine vollständige V-Kette.

Genau so ergibt sich aus Satz 20 und 21

Satz 23. *Kommen in einer Folge von Polynomen positiven Grades alle Gradzahlen $m = 1, 2, 3, \ldots$ vor, und sind alle Polynome der Folge mit einem und demselben Polynom $g(x)$ zweiten Grades vertauschbar, so gehört $g(x)$ zur Folge, und die Folge ist eine V-Kette, enthält also zu jedem Grade genau einen Vertreter.*

§ 7. Die Konstruktion aller V-Ketten.

Hat man nach Definition 4 in § 3 eine V-Kette, so enthält sie nach Satz 22 zu jeder Gradzahl $m \geq 1$ genau *einen* Vertreter und ist eine vollständige V-Folge (Definition 3 in § 3). Bezeichnet man das Element m-ten Grades der Kette mit $F_m(x)$, so ist also die Kette

$$F_1, F_2, F_3, \ldots,$$

wobei $F_1(x) \equiv x$ ist.

Für den Fall, daß alle $F_\lambda(x)$ einen gemeinsamen Eigenwert α haben, kann man die Kette durch Zufügung von $F_0(x) = \alpha$ zu einer *erweiterten Kette* machen, da ja $F_\lambda F_0 = F_0 F_\lambda = \alpha$ ist. So entstehen aus der Kette

(I) $$x, x^2, x^3, \ldots,$$

da alle Elemente in (I) sowohl 0 als auch 1 als Eigenwert besitzen, die erweiterten Ketten

(I') $$\begin{cases} 1, x, x^2, \ldots \\ 0, x, x^2, \ldots. \end{cases}$$

Transformiert man jedes F_n in einer Kette durch ein und dieselbe T-Transformation, so bilden die Polynome $T^{-1}F_nT(x)$ ebenfalls eine Kette, die zur ursprünglichen ähnlich ist. Alle zu einer gegebenen Kette ähnlichen V-Ketten sind untereinander ähnlich. Tut man alle zueinander ähnlichen V-Ketten in eine Klasse, so kann man die Frage stellen: Wie viele Klassen gibt es? Die Frage beantwortet

Satz 24. *Die Anzahl der Klassen ähnlicher V-Ketten ist gleich* 2. [Für erweiterte V-Ketten ist wegen (I') die Anzahl gleich 3.]

Beweis. Jede V-Kette enthält ein einziges Polynom zweiten Grades. Es gibt nach den im Anschluß an Satz 17 in § 5 gemachten Bemerkungen eine *einzige* T-Transformation, die dieses quadratische Polynom in ein reines Polynom x^2+b transformiert. Daher gibt es in jeder Klasse ähnlicher V-Ketten eine *einzige* Kette, in der das quadratische Polynom ein reines Polynom ist. Diese Kette soll die *Grundkette* der Klasse heißen. (I) ist die Grundkette der zu (I) gehörenden Klasse. Wir zeigen nun einfach: Es gibt nur zwei Grundketten. Damit ist dann der Satz bewiesen.

Es liege also eine Grundkette vor, in der das quadratische Polynom

$$g(x) = x^2 + b \tag{30}$$

ist. Die Grundkette enthält ein Polynom dritten Grades

$$f(x) = a_0 x^3 + a_1 x^2 + a_2 x + a_3. \tag{31}$$

Wir haben wegen der Vertauschbarkeit von $f(x)$ mit $g(x)$ die Gleichung

$$f(x^2+b) = (f(x))^2 + b.$$

Ersetzt man hier x durch $-x$, so folgt

$$(f(-x))^2 = (f(x))^2,$$

und da $f(x)$ ein Polynom ungeraden Grades ist, folgt $f(-x) = -f(x)$ [2]). Es muß also $a_1 = a_3 = 0$ sein und daher

$$f(x) = a_0 x^3 + a_2 x. \tag{31$_1$}$$

Die Vertauschbarkeitsrelation lautet somit

$$(a_0 x^3 + a_2 x)^2 + b = a_0(x^2+b)^3 + a_2(x^2+b). \tag{32}$$

Hier treten rechts und links nur gerade Potenzen von x auf. Der Vergleich der Koeffizienten von x^6 in (32) ergibt

$$a_0 = 1. \tag{33}$$

Der Vergleich der Koeffizienten von x^4 liefert

$$a_2 = \tfrac{3}{2} b. \tag{34}$$

Vergleicht man nun noch die Koeffizienten von x^2 und x^0 in (32) und trägt in die so entstehenden beiden Gleichungen die Werte für a_0 und a_1 aus (33)

und (34) ein, so erhält man nach kurzer Rechnung die beiden Gleichungen für b

$$(35) \qquad \begin{cases} b(b+2) = 0 \\ b(b+2)(2b-1) = 0, \end{cases}$$

denen b zu genügen hat, wenn

$$f(x) = x^3 + \tfrac{3}{2} b x$$

mit

$$g(x) = x^2 + b$$

vertauschbar sein soll. Die einzigen Werte b, die beide Gln. (35) erfüllen, sind

$$(36) \qquad b = 0, \quad b = -2.$$

Hiernach gibt es höchstens zwei Grundketten, die eine mit dem quadratischen Polynom x^2 und die andere mit dem Polynom zweiten Grades $x^2 - 2$. Die Grundkette mit x^2 als quadratischem Polynom ist gerade die Kette (I). Die Frage ist, ob es eine Grundkette gibt, die

$$(30_1) \qquad f_2(x) = x^2 - 2$$

als Polynom zweiten Grades besitzt. Da nach (31_1), (33) und (34)

$$(31_2') \qquad f_3(x) = x^3 - 3x$$

mit $f_2(x)$ vertauschbar ist, müßte die Grundkette so anfangen

$$f_1(x) = x, \quad f_2(x) = x^2 - 2, \quad f_3(x) = x^3 - 3x, \ldots.$$

Da ja das einzige Polynom vierten Grades, das mit f_2 vertauschbar ist (Satz 20), $f_2^2 = f_4(x) = x^4 - 4x^2 + 2$ ist, hätten wir von der vermuteten Kette schon 4 Glieder. Nun haben diese vier ersten Glieder alle den Eigenwert 2 gemeinsam. Daher ist zu vermuten, daß in der erwarteten Kette $f_0 = 2$ Eigenwert aller Kettenglieder sein wird. Wenn wir daher die Kette zu einer erweiterten Kette machen, beginnt sie:

$$f_0 = 2, \quad f_1 = x, \quad f_2 = x^2 - 2, \quad f_3 = x^3 - 3x, \quad f_4 = x^4 - 4x^2 + 2, \ldots.$$

Bei näherer Betrachtung dieser Polynome sieht man, daß sie die Rekursion

$$(37) \qquad f_n = x \cdot f_{n-1} - f_{n-2}, \qquad n = 2, 3, 4$$

erfüllen. Diese Rekursionsformel definiert sukzessive

$$f^5(x) = x^5 - 5x^3 + 5x, \quad f_6(x) = x^6 - 6x^4 + 9x^2 - 2, \ldots.$$

Wenn diese Polynome Glieder einer Kette sein sollen, muß z.B. $f_6(x) = f_2 f_3 = f_3 f_2$ sein, was man durch Rechnung bestätigt. Wenn wirklich die durch (37) rekursiv definierten $f_n(x)$ eine Kette darstellen, so zeigt (37) auch, daß alle Polynome der Kette den gemeinsamen Eigenwert 2 besitzen.

Definiert man noch

$$f_{-1} = x, \tag{37$_1$}$$

so gilt (37) schon von $n = 1$ an.

Nach Satz 23 brauchen wir zum Nachweise, daß die durch (37) definierten Polynome eine V-Kette bilden, nur zu zeigen, daß jedes $f_n(x)$ mit $f_2(x)$ vertauschbar ist. Es ist also

$$f_n(x^2 - 2) = (f_n(x))^2 - 2 \tag{38}$$

für alle n zu zeigen. Für $n = -1, 0, 2, 3, 4$ ist ja (38) richtig. Die Richtigkeit von (38) wird durch Induktion bewiesen. Dazu bemerken wir, daß der Ausdruck

$$(f_{n-1}(x))^2 + (f_n(x))^2 - x f_n(x) \cdot f_{n-1}(x)$$

vom Index n unabhängig ist. Denn ersetzt man in ihm nach (37) $f_n(x)$ durch $x f_{n-1} - f_{n-2}$, so geht er in den Ausdruck

$$(f_{n-2}(x))^2 + (f_{n-1}(x))^2 - x f_{n-1}(x) \cdot f_{n-2}(x)$$

über. Da der Ausdruck für $n = 1$ gleich $4 - x^2$ ist, gilt also

$$(f_{n-1})^2 + (f_n)^2 - x \cdot (f_n) \cdot (f_{n-1}) = 4 - x^2. \tag{37$_2$}$$

Nehmen wir nun an, daß (38) schon richtig ist für $r = 0, 1, 2, \ldots, n$:

$$f_r(x^2 - 2) = (f_r(x))^2 - 2, \tag{38$_1$}$$

so zeigen wir mit Hilfe von (37$_2$), daß (38$_1$) auch für $r = n + 1$ gilt. Dazu ersetzen wir in (37) x durch $x^2 - 2$ und erhalten

$$f_{n+1}(x^2 - 2) = (x^2 - 2) f_n(x^2 - 2) - f_{n-1}(x^2 - 2),$$

wenn wir (37) für den Index $(n + 1)$ benutzen. Wendet man rechts die schon für $r \leq n$ als richtig angenommene Gl. (38$_1$) an, so erhält man nach kurzer Rechnung:

$$f_{n+1}(x^2 - 2) = (x \cdot f_n)^2 - 2\{x^2 + (f_{n-1})^2 + (f_n)^2\} + (f_{n-1})^2 + 6.$$

Da die geschweifte Klammer auf der rechten Seite nach (37$_2$) gleich

$$4 + x \cdot f_{n-1} \cdot f_n$$

ist, folgt unter nochmaliger Verwendung von (37)

$$f_{n+1}(x^2 - 2) = (f_{n+1}(x))^2 - 2,$$

womit die Richtigkeit von (38$_1$) auch für $r = n + 1$ gezeigt ist. Also gilt (38) stets und damit gibt (37) wirklich eine Grundkette

$$\text{(II)} \qquad \begin{cases} f_1 = x, \quad f_2 = x^2 - 2, \quad f_3 = x^3 - 3x, \ldots \\ f_{n+1} = x f_n - f_{n-1} \quad (n = 0, 1, 2, \ldots), \quad f_0 = 2, \quad f_{-1} = x. \end{cases}$$

Alle Polynome in (II) haben den Eigenwert 2. Jede V-Kette ist also entweder zu (I) oder zu (II) ähnlich. (I) und (II) sind in den beiden Klassen die Grundketten.

Wie in jeder Kette, so ist auch hier

$$f_m f_n = f_n f_m = f_{mn}. \tag{39}$$

Die Polynome der Kette (II) sind aus der Algebra wohlbekannt, und von dorther ergibt sich auch unmittelbar die Gl. (39). Setzt man nämlich für variables z

$$z + \frac{1}{z} = x, \quad s_n = z^n + z^{-n}, \quad n = 0, \pm 1, \pm 2, \ldots, \tag{40}$$

so ist $s_1 = x$ und

$$s_n s_1 = s_n x = s_{n+1} + s_{n-1},$$

$$s_{n+1} = x s_n - s_{n-1}, \quad s_0 = 2, \quad s_{-1} = s_1 = x.$$

Daher wegen (37)

$$s_n = f_n(x) = f_n\left(z + \frac{1}{z}\right) = z^n + z^{-n}. \tag{41}$$

Ersetzt man hier z durch z^m, so folgt sofort

$$\begin{aligned} z^{nm} + z^{-nm} = s_{mn} = f_{mn}(x) &= f_n(z^m + z^{-m}) \\ &= f_n f_m(x) \\ &= f_m f_n(x), \end{aligned}$$

womit alles von neuem gezeigt ist.

Nun steht schon bei J. F. RITT[1]) als Beispiel für vertauschbare Polynome das Beispiel der Tschebyscheffschen Polynome, wobei RITT die Vertauschbarkeit aus der Definitionsgleichung

$$\cos(n\varphi) = T_n(\cos\varphi), \qquad n = 0, 1, 2, \ldots \tag{42}$$

erhält, indem er φ durch $m\varphi$ ersetzt. Diese Ritt-Tschebyscheffsche V-Kette ist also

$$T_1 = x, \quad T_2 = 2x^2 - 1, \ldots. \tag{III}$$

Diese Kette ist zu (II) ähnlich und geht aus (II) durch die Ähnlichkeitstransformation $T(2, 0)$ hervor, so daß

$$\frac{1}{2} f_n(2x) = T_n(x), \qquad f_n(x) = 2\, T_n\left(\frac{x}{2}\right) \tag{43}$$

ist. Man erhält dies ja auch unmittelbar, indem man oben $z = e^{i\varphi}$, $x = 2\cos\varphi$

$$s_n = f_n(x) = f_n(2\cos\varphi) = 2\cos(n\varphi)$$

einführt. Aus (37) ergibt sich dann die Rekursionsformel für die Tschebyscheffschen Polynome.

Für die Polynome $f_n(x)$ hat man die explizite Darstellung[3])

$$f_n(x) = \sum_{k=0}^{\left[\frac{n}{2}\right]} \left\{2\binom{n-k}{k} - \binom{n-k-1}{k}\right\} (-1)^k\, x^{n-2k}, \qquad n = 1, 2, 3, \ldots. \tag{37$_3$}$$

[3]) Siehe JACOBSTHAL, E.: Über die Klasseninvariante linearer Abbildungen. II. Det Kongelige Norske Videnskab. Selsk. Forh. **26**, 10—15, (1953) [Gl. (15a) auf S. 12].

Es sei bemerkt, daß aus der Gleichung

$$\cos^2(n\varphi) = \psi_n(\cos^2\varphi)$$

Polynome definiert werden, die ebenfalls eine Kette der Klasse (II) bilden. Die Kette der $\psi_n(x)$ geht durch $T(\frac{1}{2}, \frac{1}{2})$ in die Ritt-Tschebyscheffsche Kette über. — Es soll noch auf eine bemerkenswerte Tatsache hingewiesen werden: Ist ein quadratisches Polynom mit einem kubischen Polynom vertauschbar, dann besitzen sie beide vertauschbare Polynome jeden Grades. Diese Tatsache wird durch Satz 35 wesentlich verschärft werden.

§ 8. Ein zahlentheoretischer Satz über vertauschbare Polynome.

Ist $g(x)$ ein Polynom eines Grades ≥ 2 mit algebraischen Zahlkoeffizienten, und ist $f(x)$ mit $g(x)$ vertauschbar, dann folgt aus § 6 leicht, daß die Koeffizienten von $f(x)$ auch algebraische Zahlen sind. Nun betrachten wir speziell solche Polynome $g(x)$, bei denen alle Koeffizienten *ganze* algebraische Zahlen sind, deren höchster sogar eine *Einheit* sein soll. Solche Polynome $g(x)$ wollen wir kurz G-Polynome nennen. Ihre Nullstellen sind bekanntlich ganze algebraische Zahlen. Ist der Grad eines G-Polynoms $g(x)$ größer als 1, so sind daher die Eigenwerte ganze Zahlen. Die Iterierten eines G-Polynoms sind ebenfalls G-Polynome. Wir haben später Verwendung für einen Satz über G-Polynome, der folgendermaßen lautet:

Satz 25. *Ist $g(x)$ ein G-Polynom eines Grades >1, so ist jedes mit $g(x)$ vertauschbare Polynom $f(x)$ positiven Grades ebenfalls ein G-Polynom.*

Beweis. Es sei

$$g(x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_n, \qquad n \geq 2,$$

worin alle b_k ganze algebraische Zahlen sind und b_0 eine Einheit ist. Ferner sei

$$f(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_m, \qquad m \geq 1,$$

mit $g(x)$ vertauschbar. Nach (28_1) in § 6 ist dann

$$a_0^{n-1} = b_0^{m-1},$$

also a_0 auch eine Einheit. Der Fall, daß $f(x) \equiv x$ ist, kann außer Betracht bleiben, da ja dann $f(x)$ ein G-Polynom ist. Der Fall $f(x) = x + b$, $b \neq 0$, ist nach Satz 9 unmöglich, da der Grad n von $g(x)$ größer als 1 ist. Die einzigen linearen Polynome $f(x)$, die in Frage kommen können, sind

$$f(x) = a_0 x + a_1,$$

worin a_0 eine von 1 verschiedene Einheit ist. In diesem Falle besitzt $f(x)$ einen einzigen Eigenwert. Jedes mit $g(x)$ vertauschbare Polynom $f(x) \not\equiv x$ hat also m Eigenwerte, worin m der Grad von $f(x)$ ist.

Sei ζ ein beliebiger Eigenwert von $f(x)$. Da $g(x)$ mit $f(x)$ vertauschbar ist, ist auch jede Iterierte g^p von g mit $f(x)$ vertauschbar; also ist nach Satz 11 jede der Zahlen $g^p(\zeta)$ auch ein Eigenwert von $f(x)$; deswegen enthält die Folge

$$g^0(\zeta)=\zeta,\quad g(\zeta),\quad g^2(\zeta),\ldots$$

nur endlich viele verschiedene Zahlen, da ja $f(x)$ nur m Eigenwerte hat. Es ist also für passende r, s, $r\geq 0$, $s\geq 1$

$$g^r(\zeta)=g^{r+s}(\zeta),$$

und dies ist eine algebraische Gleichung für ζ, deren Koeffizienten ganze algebraische Zahlen sind, deren höchster eine Einheit ist, da ja die Iterierte eines G-Polynomes wieder ein G-Polynom ist. Also ist ζ eine ganze algebraische Zahl. Alle Eigenwerte von $f(x)$ sind also ganze algebraische Zahlen. Ist nun $m=1$, $f(x)=a_0x+a_1$, so ist

$$\zeta=\frac{a_1}{1-a_0}$$

eine ganze algebraische Zahl, also $a_1=\zeta(1-a_0)$ auch eine solche und $f(x)$ ein G-Polynom. Ist aber $m>1$, und sind die m Eigenwerte von $f(x)$ etwa $\zeta_1,\zeta_2,\ldots,\zeta_m$, dann ist wegen $m>1$ der höchste Koeffizient in $f(x)-x$ auch a_0 und daher

$$f(x)-x=a_0(x-\zeta_1)\ldots(x-\zeta_m),$$

$$f(x)=a_0(x-\zeta_1)\ldots(x-\zeta_m)+x,$$

worin a_0 eine Einheit und alle ζ_k ganze algebraische Zahlen sind. Also ist $f(x)$ ein G-Polynom, w. z. b. w.

§ 9. Hilfsbetrachtungen über eine Zahlfolge.

Die Frage, wann bei gegebenem $g(x)$ die Gl. (1) als Lösungen $f(x)$ nur die Iterierten von $g(x)$ besitzt, oder ob es noch andere Lösungspolynome $f(x)$ gibt, soll später für das reine Polynom $g(x)=x^r+a$, $(r=2,3,\ldots)$ und für alle quadratischen Polynome entschieden werden. Für die Beantwortung dieser Frage ist es von Wichtigkeit zu wissen, ob die Zahl 0 eine reguläre oder singuläre Zahl des Polynoms $g(x)=x^r+a$ ist. Man hat also die Zahlfolge

$$V_{g(0)}=\{g^{(n)}(0)\}_{n=0,1,2,\ldots}$$

zu betrachten. Unter Weglassung der ersten Zahl sind das die Zahlen

$$a_n=g^{(n+1)}(0),\qquad n=0,1,2,\ldots. \tag{44}$$

Es ist also

$$\begin{cases} a_0=g(0)=a,\\ a_1=g(a)=a^r+a,\\ a_2=g(a_1)=a_1^r+a=(a^r+a)^r+a,\\ \quad\vdots \end{cases} \tag{44$_1$}$$

$$a_{n+1}=a_n^r+a,\quad a_0=a,\qquad n=0,1,2,\ldots. \tag{44$_2$}$$

Die a_n sind ganzzahlige Polynome von a,

$$(45)\qquad \begin{cases} a_n = \psi_n(a), \\ \psi_{n+1}(a) = (\psi_n(a))^r + a, \quad \psi_0(a) = a. \end{cases}$$

Das niederste Glied in jedem dieser Polynome ist a, der höchste Koeffizient ist gleich 1. Die Gleichung $\psi_n(a) = 0$ hat also außer $a = 0$ nur Einheiten als Lösung, falls $n \geq 1$ ist. Bei $n \geq 1$ sind alle Glieder außer dem letzten durch a^r teilbar. Die Frage, ob 0 eine reguläre oder singuläre Zahl des Polynoms $g(x) = x_r + a$ ist, ist gleichbedeutend mit der Frage: Für welche a sind alle a_n verschieden und für welche a ist die Folge $\{a_n\}$ periodisch? Das letztere tritt dann und nur dann ein, wenn für zwei verschiedene Indizes l, m, $l < m$,

$$(46)\qquad a_m = a_l,$$

$$(46_1)\qquad \psi_m(a) = \psi_l(a)$$

ist. Ist dann m der kleinste Index, für den ein a_m gleich einem a_l mit $l < m$ wird, so sind die Zahlen $a_0, a_1, \ldots, a_{m-1}$ alle verschieden. Die ganze Folge enthält nur m verschiedene Zahlen; von a_l an wiederholen sich die Zahlen der Folge periodisch. Wenn $l \geq 1$ ist, bilden die Zahlen $a_0, a_1, \ldots, a_{l-1}$ eine Vorperiode, deren Glieder untereinander und von allen folgenden Zahlen verschieden sind. Für $l = 0$ fehlt die Vorperiode und die Folge ist reinperiodisch. Da die Gln. (46), (46_1) algebraische Gleichungen mit ganzen Koeffizienten sind, deren höchster gleich 1 ist, wird die Folge (44) periodisch, wenn a eine ganze algebraische Zahl ist, die (46), (46_1) genügt. Diese Zahlen wollen wir *kritische* Zahlen für den Exponenten r nennen; die anderen Zahlen a sollen *gewöhnliche* Zahlen für den Exponenten r heißen. Für eine gewöhnliche Zahl a ist also 0 eine reguläre Zahl des Polynoms $x^r + a$, dagegen ist 0 eine singuläre Zahl des Polynoms, falls a eine kritische Zahl für den Exponenten r ist. Alle Lösungen der Gln. (46), (46_1) stellen die Gesamtheit der kritischen Zahlen dar. Unter ihnen befindet sich trivialerweise die Zahl 0, da $\psi_n(0) = 0$ für $n = 0, 1, 2, \ldots$ ist. Diese können wir später außer Betracht lassen. Für den Exponenten $r = 2$ sind z.B. einige kritische Zahlen a und die zugehörigen $\{a_n\}$ in der Tabelle angegeben:

$a_0 = a$	a_1	a_2	a_3	a_4	a_5	...
0	0	0	0	0	0	...
-1	0	-1	0	-1	0	...
-2	2	2	2	2	2	...
i	$-1+i$	$-i$	$-1+i$	$-i$	$-1+i$	...
$-i$	$-1-i$	i	$-1-i$	i	$-1-i$	...
⋮	⋮	⋮	⋮	⋮	⋮	

Wir beweisen nun über die reellen kritischen Zahlen

Satz 26. *Für ungerades $r>2$ ist die einzige reelle kritische Zahl $a=0$. Für gerades $r\geq 2$ sind $a=0$, $a=-1$ und $a=-\sqrt[r-1]{2}$ kritische Zahlen. Alle sonst noch etwa vorhandenen reellen kritischen Zahlen liegen zwischen $-\sqrt[r-1]{2}$ und -1.*

Beweis. Für $a>0$ ist ja

$$a_0<a_1<a_2<\cdots.$$

Also ist jede positive Zahl eine gewöhnliche Zahl. Da für ungerades r

$$\psi_n(-a)=-\psi_n(a)$$

ist, so folgt für ungerade r, daß auch jede negative Zahl a eine gewöhnliche Zahl ist. Daher ist $a=0$ die einzige reelle kritische Zahl für den ungeraden Exponenten r.

Sei nun $r=2, 4, 6, \ldots$ und $a<0$.

Für $a=-1$ sieht dann die Folge so aus wie in der Tabelle des Falles $r=2$, so daß -1 kritische Zahl ist. Für $a_0=a=-\sqrt[r-1]{2}$ wird $a_1=a_2=a_3=\cdots=-a_0=\sqrt[r-1]{2}$, so daß $-\sqrt[r-1]{2}$ kritische Zahl ist. (Siehe Tabelle für $r=2$.)

Für $a<-\sqrt[r-1]{2}$ ist $-a>\sqrt[r-1]{2}$ und daraus folgt der Reihe nach

$$-a^{r-1}>2,\quad -a^r<2a,\quad a_1=a^r+a>-a>0,\quad a_1>|a|,\quad a_1^r>a_0^r=a^r,$$

$a_2>a_1$ und daher

$$0<a_1<a_2<a_3<\cdots.$$

Also ist jedes $a<-\sqrt[r-1]{2}$ eine gewöhnliche Zahl. Ist nun schließlich $-1<a<0$, so zeigt man auf ähnlichem Wege wie eben, daß

$$-1<a_0<a_2<a_4<\cdots<a_5<a_3<a_1<0$$

ist, also sind alle a_n voneinander verschieden und a keine kritische Zahl. Wenn es also außer den drei genannten reellen kritischen Zahlen noch andere solche geben sollte, liegen sie zwischen $-\sqrt[r-1]{2}$ und -1, w. z. b. w.

Wir beweisen nun weiter

Satz 27. *Ist die kritische Zahl $a\neq 0$ so beschaffen, daß die zugehörige periodische Folge $\{a_n\}$ keine Vorperiode hat, also reinperiodisch ist, dann ist a eine algebraische Einheit. Ist dagegen eine Vorperiode vorhanden und existiert eine Primzahl q, die in einer Potenz a^λ, $1\leq\lambda\leq r-1$, aufgeht, dann ist notwendigerweise $r=q=2$ und a ist eine zu 2 assoziierte algebraische Zahl.*

Beweis. Die kritischen Zahlen sind ganze algebraische Zahlen. Sei $a\neq 0$ eine solche und m, l, $m>l$, die Minimalindizes, für welche die a definierenden Gln. (46), (46_1) gelten. Ist dann $\{a_n\}$ reinperiodisch, so ist $l=0$, $a_m=a_{m-1}^r+a=a_0=a$; also ist $a_{m-1}=0$ und wegen $a_0=a\neq 0$ ist $m-1>0$; $m\geq 2$. Hieraus folgt:

$$a_{m-1}=0=a_{m-2}^r+a.$$

Das ist nach Division mit a eine algebraische Gleichung mit ganzen Koeffizienten, deren höchster und niederster gleich 1 ist. Also ist a eine Einheit.

Damit ist der erste Teil erledigt. — Nun sei a eine solche kritische Zahl $\neq 0$, daß $\{a_n\}$ eine Vorperiode hat. Es ist $l \geq 1$. Aus (46) folgt dann

$$a_{m-1}^r = a_{l-1}^r,$$

$$a_{m-1} = \varepsilon a_{l-1},$$

worin ε eine r-te Einheitswurzel ist, die wegen der Minimaleigenschaft von m von 1 verschieden ist. Es sei ε eine primitive s-te Einheitswurzel, $s \geq 2$, $r \equiv 0 \bmod s$. Da jedes a_n als tiefstes Glied das Glied a hat und alle vor ihm stehenden Glieder Exponenten $\geq r$ besitzen, ergibt die Gleichung $a_{m-1} - \varepsilon a_{l-1} = 0$ nach Division mit a, daß

(a) $$1 - \varepsilon \equiv 0 \bmod a^{r-1}$$

ist. Falls es nun eine Primzahl q gibt, die in einer Potenz a^λ aufgeht, worin $1 \leq \lambda \leq r-1$ ist, so sei

(b) $$\left[\frac{r-1}{\lambda}\right] = \mu, \quad 1 \leq \lambda\mu \leq r-1.$$

Dann ist

(a_1) $$1 - \varepsilon \equiv 0 \,(q^{\lambda\mu}).$$

In dieser Kongruenz kann ε durch jede der $\varphi(s)$ konjugierten ε^t, $1 \leq t \leq s-1$, $(t, s) = 1$, ersetzt werden. Die Multiplikation der $\varphi(s)$ Kongruenzen (a_1) ergibt dann

(c) $$F_s(1) \equiv 0 \bmod q^{\lambda\mu\varphi(s)},$$

worin $F_s(x)$ das s-te Kreisteilungspolynom ist. Da nun $F_s(1) = 1$ ist, falls s durch zwei verschiedene Primzahlen teilbar ist, und $F_s(1) = p$ ist, falls s die Potenz einer Primzahl p ist, folgt, daß wegen (c)

$$s = q^\alpha$$

sein muß und

$$q \equiv 0 \bmod q^{\lambda\mu\varphi(s)}.$$

Also folgt

$$\lambda\mu\varphi(s) = 1, \quad \lambda = 1, \quad \mu = 1, \quad \varphi(s) = 1, \quad s = 2 = q^\alpha, \quad q = 2, \quad \alpha = 1.$$

Wegen (b) ist $r - 1 = 1$, $r = 2$. Da $s = 2$ ist, wird $\varepsilon = -1$ und daher nach (a) die ganze algebraische Zahl a ein Teiler von $1 - \varepsilon = 2$. Es war aber $q = 2$ nach Annahme ein Teiler von $a^\lambda = a^1 = a$. Also ist 2 zu a assoziiert und der Satz bewiesen. Für den Exponenten $r = 2$ ist ja die kritische Zahl -2 zu 2 assoziiert. Wir beweisen nun

Satz 28. *Für den Exponenten $r = 2$ ist -2 die einzige zu 2 assoziierte kritische Zahl.*

Beweis. Die a_n sind bei $r = 2$ durch

$$a_{n+1} = a_n^2 + a, \quad a_0 = a,$$

definiert:

$$a_0 = a, \quad a_1 = a^2 + a, \quad a_2 = (a^2 + a)^2 + a, \ldots.$$

Rechnet man hier, um die kritischen Zahlen a zu erhalten, die niedersten Fälle der Gleichung $a_m = a_l$, $m > l$, durch, also die Fälle:

$$a_1 = a; \quad a_2 = a; \quad a_3 = a;$$
$$a_2 = a_1; \quad a_3 = a_1;$$
$$a_3 = a_2,$$

so erhält man als kritische Zahlen:

$$0, \ -1, \ -2, \ +i, \ -i;$$

außerdem noch aus $a_3 = a$ die Lösungen von

$$a^3 + 2a^2 + a + 1 = 0,$$

die Einheiten sind, und im Falle $a_3 = a_2$ noch die Gleichung

$$a^3 + 2a^2 + 2a + 2 = 0.$$

Aus ihr folgt, daß $a \mid 2$; $a^2 \mid 2$; $a^3 \mid 2$ ist. Also ist

$$2 = a^3 \eta,$$

worin η eine ganze algebraische Zahl ist. Also bekommt man

$$a^3 + a^5 \eta + a^4 \eta + a^3 \eta = 0$$
$$1 + \eta (a^2 + a + 1) = 0.$$

Daher ist η eine Einheit. Da a^3 keine Einheit ist, ist a^2 und $a^2\eta$ keine Einheit, also auch $2/a$ keine Einheit, a also nicht zu 2 assoziiert.

In den niedersten Fällen ist also die einzige kritische zu 2 assoziierte Zahl -2.

Es ist daher für $N = 4$ richtig, daß unter allen kritischen Zahlen, die man aus

$$a_m = a_l, \quad l < m < N = 4$$

erhält, -2 die einzige zu 2 assoziierte Zahl ist. Wir nehmen nun an, es sei für eine positive ganze Zahl N richtig, daß unter allen kritischen Zahlen a, die sich aus allen Gleichungen

$$a_m = a_l, \quad 0 \leq l < m < N$$

ergeben, -2 die einzige zu 2 assoziierte Zahl ist. Wenn wir dann zeigen können, daß die N Gleichungen $a_N = a_l$, $l = 0, 1, \ldots, N-1$ außer -2 keine zu 2 assoziierte Zahl liefern, dann ist die Richtigkeit der Annahme auf alle Gleichungen

$$a_m = a_l, \quad 0 \leq l < m < N + 1$$

ausgedehnt, und da der Satz für $N = 2, 3, 4$ richtig war, gilt er dann allgemein.

Von den N Gleichungen $a_N = a_l$ ist die erste $a_N = a$, die $a_{N-1} = 0$ ergibt. Das ergibt außer $a = 0$ nur Einheiten a, also keine zu 2 assoziierte Zahl. Sei also $1 \leq l \leq N-1$. Dann ergibt die Gleichung $a_N = a_l$, daß $a_{N-1}^2 = a_{l-1}^2$ ist. Also wird $a_{N-1} = a_{l-1}$ oder $a_{N-1} = -a_{l-1}$. Von der ersten dieser Gleichungen stammen nach Annahme nur kritische Zahlen, die nicht zu 2 assoziiert sind, außer etwa -2. Also brauchen wir nur noch die zweite Gleichung

$$a_{N-1} + a_{l-1} = 0$$

zu betrachten. Für $l=1$ ergibt sich die Gleichung

$$a_{N-1}+a_0=a_{N-1}+a=0$$

oder

$$\psi_{N-1}(a)+a=0.$$

Nun ist das Polynom $\psi_{N-1}(x)+x$ gleich 0 für $x=0$ und $x=-2$. Also ist $\psi_{N-1}(x)+x$ durch $x(x+2)$ teilbar. Da $N\geq 4$, $N-1\geq 3$ ist, sind x^2+x die beiden niedersten Glieder in $\psi_{N-1}(x)$, daher sind die beiden niedersten Glieder in $\psi_{N-1}(x)+x$ gerade $x^2+x+x=x(x+2)$. Wenn man also $\psi_{N-1}(x)+x$ durch $x(x+2)$ dividiert, erhält man ein Polynom mit ganzen Koeffizienten, von denen der höchste und tiefste gleich 1 ist. Also hat die Gleichung $\psi_{N-1}(a)+a=0$ als Lösungen nur $a=0$, $a=-2$ und außerdem nur Einheiten. Nun sei $l\geq 2$. Da sind a^2 und a die niedersten Glieder in a_{N-1} und a_{l-1}. Daher endet $a_{N-1}+a_{l-1}$ mit den beiden Gliedern $2a^2+2a$. Nach Abspaltung des Faktors a endet also das übrigbleibende Polynom mit

$$\cdots+2a+2.$$

Die nach Abspaltung des Faktors a aus $a_{N-1}+a_{l-1}=0$ entstehende Gleichung

$$\cdots+2a+2=0$$

zeigt also, daß 2 durch die kritische Zahl a und daher sogar durch a^2 teilbar wird. Es ist also $2=a^2\eta$, worin η eine ganze algebraische Zahl ist. Da 2 keine Einheit ist, können nicht gleichzeitig a und η Einheiten sein, also ist auch $a\eta=2/a$ keine Einheit, also ist die kritische Zahl a nicht zu 2 assoziiert, womit durch Induktionsschluß der Satz bewiesen ist.

Aus Satz 27 und 28 folgt nun

Satz 29. *Ist $a\neq 0$ eine kritische Zahl für den Exponenten r und ist a durch eine Primzahl teilbar, dann ist $r=2$ und $a=-2$.*

Beweis. Ist nämlich a durch eine Primzahl q teilbar, dann ist a keine Einheit; die zu a gehörige Folge $\{a_n\}$ hat nach Satz 27 eine Vorperiode und nach demselben Satz ist $r=2$ und a zu 2 assoziiert. Nach Satz 28 aber ist bei $r=2$ die einzige zu 2 assoziierte kritische Zahl $a=-2$, womit alles gezeigt ist.

§ 10. Das Polynom $g(x)=x^r+a$.

Wir setzen voraus, daß der Grad $r\geq 2$ und $a\neq 0$ ist, und beweisen

Satz 30. *Jedes mit $g(x)$ vertauschbare Polynom $f(x)$ des Grades $n>0$ hat die Gestalt*

$$f(x)=c_n x^n+c_{n-r}x^{n-r}+c_{n-2r}x^{n-2r}+\cdots \quad (47)$$

[4]) Dieser Satz gilt allgemein, wenn $g(x)$ nur Potenzen von x^r enthält. Dann ist für eine primitive r-te Einheitswurzel ε stets $g(\varepsilon x)=g(x)$, und wenn $fg=gf$ ist, folgt $gf(x)=fg(x)=fg(\varepsilon x)=gf(\varepsilon x)$. Aus $g(f(\varepsilon x))-g(f(x))=0$ läßt sich dann der Faktor $(f(\varepsilon x))^r-((fx))^r$ abspalten, und die Untersuchung des höchsten Gliedes des Restfaktors zeigt, daß er $\neq 0$ ist, so daß man (49) erhält, woraus nach dem Beweis im Text alles Weitere folgt.

Beweis. Da $fg = gf$ ist, gilt

$$f(x^r + a) = (f(x))^r + a. \tag{48}$$

Bedeutet ε eine primitive r-te Einheitswurzel, so bleibt die linke Seite von (48) unverändert, wenn man x durch εx ersetzt. Da also auch die rechte Seite sich nicht ändert, folgt

$$(f(\varepsilon x))^r = (f(x))^r. \tag{49}$$

Daher wird

$$f(\varepsilon x) = \varepsilon^c (f(x)), \tag{50}$$

worin c eine passende ganze Zahl ist. Ist nun

$$f(x) = \sum_{\nu=0}^{n} c_\nu x^\nu, \quad c_n \neq 0, \tag{51}$$

so folgt aus (50) durch Koeffizientenvergleichung

$$c_\nu \varepsilon^\nu = \varepsilon^c c_\nu, \quad \nu = 0, 1, \ldots, n. \tag{52}$$

Insbesondere für $\nu = n$ wegen $c_n \neq 0$

$$\varepsilon^n = \varepsilon^c.$$

Daher ist

$$c_\nu(\varepsilon^n - \varepsilon^\nu) = 0. \tag{52_1}$$

Hieraus folgt, daß alle Koeffizienten $c_\nu = 0$ werden, bei denen $n - \nu \not\equiv 0 \bmod r$ ist. Das Polynom $f(x)$ enthält also nur solche Glieder $c_\nu x^\nu$, bei denen $n - \nu = \lambda r$, $\nu = n - \lambda r$ ist, w. z. b. w.

Nun teilen wir alle mit $g(x^r) = x + a$ vertauschbaren Polynome positiven Grades in zwei Abteilungen ein:

Abteilung A enthält alle diejenigen $f(x)$, die einen durch r teilbaren Grad n haben. — Für jedes $f(x)$ in A ist der Grad $\geq r$.

Abteilung B enthält alle übrigen mit $g(x)$ vertauschbaren Polynome $f(x)$.

Zu A gehört $g(x)$ und alle Iterierten g^q, $q \geq 1$. Zu B gehört $f(x) \equiv x$. Es kann sein, daß B nur aus diesem Polynom besteht.

Satz 31. *In B ist $f(x) \equiv x$ das einzige Polynom, das einen Grad $n < r$ besitzt.*

Beweis. Ist nämlich für ein Polynom $f(x)$, das mit $g(x)$ vertauschbar ist, der Grad $n < r$, so ist wegen (47) $f(x)$ von der Gestalt

$$f(x) = c_n x^n.$$

Daher erhalten wir aus (48) die Beziehung

$$c_n(x^r + a)^n = c_n^r x^{nr} + a, \tag{53}$$

die nur für $n = 1$ möglich ist und dann $c_n = 1$, also $f(x) \equiv x$ ergibt.

Satz 32. *Jedes Polynom $f(x)$ der Abteilung A läßt sich zusammensetzen aus einer Iterierten g^q, $q \geq 1$, und einem Polynom $F(x)$ der Abteilung B.*

Beweis. $f(x)$ habe den Grad n, der durch r teilbar ist. Wir nehmen an, daß die höchste Potenz von r, die in n aufgeht, die q-te Potenz ist, und setzen

$$n = r^q n'; \quad n' \not\equiv 0 \bmod r. \tag{54}$$

Wir bezeichnen $f(x)$ mit $f_0(x)$, dann treten in $f_0(x)$ wegen (47) nur solche Glieder auf, deren Exponenten Multipla von r sind; also ist

$$f_0(\sqrt[r]{x-a}) = f(\sqrt[r]{x-a}) = f_1(x) \tag{55}$$

ein Polynom von x mit dem höchsten Koeffizienten c_n und dem Grade

$$n_1 = \frac{n}{r} = r^{q-1} n'; \quad n = n_0; \quad n_1 = \frac{n_0}{r}. \tag{56}$$

Dieses neue Polynom $f_1(x)$ ist ebenfalls mit $g(x)$ vertauschbar, denn es ist

$$f_1 g = f(\sqrt[r]{g(x) - a}) = f(x) = f_0(x),$$

da $g(x) - a = x^r$ ist. Und es wird

$$\begin{aligned} g f_1 &= g f(\sqrt[r]{x-a}) = f g(\sqrt[r]{x-a}) \\ &= f(x - a + a) = f(x) = f_0(x). \end{aligned}$$

$$f_1 g = g f_1 = f_0 = f. \tag{57}$$

Ist nun $q = 1$, so ist $n_1 = n'$, also der Grad von f_1 nicht durch r teilbar. Ist aber $q > 1$, so ist der Grad von f_1 noch durch r teilbar und f_1 hat, da es mit g vertauschbar ist, nur solche Glieder, deren Exponenten Multipla von r sind. Dann wird nach demselben Schlußverfahren

$$f_1(\sqrt[r]{x-a}) = f_2(x) \tag{55_1}$$

ein mit g vertauschbares Polynom des Grades

$$n_2 = \frac{n_1}{r} = r^{q-2} n, \tag{56_1}$$

und analog zu (57)

$$f_2 g = g f_2 = f_1. \tag{57_1}$$

Bei $q = 2$ sind wir am Ende des Verfahrens, bei $q > 2$ setzen wir es fort, bis zum ersten Male in der Reihe der abnehmenden Grade $n = n_0, n_1, n_2, \ldots$ ein durch r unteilbarer Grad $(= n')$ auftaucht. Wir erhalten also der Reihe nach die Polynome $f_0, f_1, f_2, \ldots, f_q$ mit den Gradzahlen

$$n_0, n_1, n_2, \ldots, n_q = n'$$

durch die Formeln

$$f_\lambda(\sqrt[r]{x-a}) = f_{\lambda-1}(x), \quad \lambda = 0, 1, \ldots, q-1. \tag{55'}$$

Der Grad von $f_\lambda(x)$ ist

$$n_\lambda = \frac{n_0}{r^\lambda} = r^{q-\lambda} n', \quad \lambda = 0, 1, 2, \ldots, q. \tag{56'}$$

$$f_\lambda g = g f_\lambda = f_{\lambda-1}, \quad \lambda = 1, 2, \ldots, q. \tag{57'}$$

Alle f_λ haben denselben höchsten Koeffizienten c_n und

$$F(x) = f_q(x) \tag{58}$$

hat den durch r unteilbaren Grad $n_q = n'$. Aus (57') erhält man durch sukzessive Anwendung ($\lambda = 1, 2, \ldots, q$):

$$f_0 = f = g f_1 = g^2 f_2 = g^3 f_3 = \cdots = g^q f_q = g^q F = F g^q$$

$$f(x) = F g^q = g^q F. \tag{59}$$

Da der Aufbau der Polynome $f(x)$ in A nach Satz 31 auf Polynome $F(x)$ in B führt, betrachten wir nun nur noch Polynome $f(x)$ der Abteilung B. Enthält die Abteilung B außer $f(x) \equiv x$ noch andere Polynome, so sei $f(x)$ eines von diesen, also nach Satz 30: $n \geq r$, $n \not\equiv 0 \bmod r$.

Satz 33. *Hat das Polynom $f(x)$ der Abteilung B einen Grad $n > r$, so ist notwendigerweise $r = 2$ und $a = -2$, d.h. wenn $g(x) \equiv x^2 - 2$ ist, besteht die Abteilung B nur aus dem Polynom $f(x) = x$.*

Beweis. Da $n \not\equiv 0 \bmod r$ ist, folgt aus (47), daß $f(x)$ kein konstantes Glied besitzt; also ist $f(0) = 0$ und somit 0 ein Eigenwert des Polynoms $f(x)$. Nun ist jede Iterierte von $g(x)$ mit $f(x)$ vertauschbar. Und deswegen ist nach Satz 11 jede der Zahlen $g^m(0)$ ein Eigenwert von $f(x)$; da aber $f(x)$ nur n Eigenwerte hat, enthält die Folge der Zahlen $a_m = g^{m+1}(0)$ nur endlich viele verschiedene, und somit ist a eine kritische Zahl für den Exponenten r. Da eine solche aber eine ganze algebraische Zahl ist, ist $g(x) = x^r + a$ ein G-Polynom und deswegen nach Satz 25 auch $f(x)$ ein solches. Daher sind die Koeffizienten $c_n, c_{n-r}, c_{n-2r}, \ldots$ in (47) ganze algebraische Zahlen, c_n eine Einheit, die nach (28_1) eine $(r-1)$-te Einheitswurzel sein muß, was sich ja aus $fg = gf$ oder

$$c_n (x^r + a)^n + c_{n-r} (x^r + a)^{n-r} + \cdots = (c_n x^n + c_{n-r} x^{n-r} + \cdots)^r + a$$

durch Koeffizientenvergleich der höchsten Glieder ergibt. Das zweithöchste Glied hat den Exponenten $r(n-1)$. Der Koeffizient von $x^{r(n-1)}$ auf der linken Seite ist $n c_n a$ und auf der rechten Seite $r c_n^{r-1} c_{n-r} = r c_{n-r}$. Daher erhalten wir

$$c_{n-r} = \frac{n a}{r} c_n. \tag{60}$$

Hierin ist c_n eine Einheit, c_{n-r} eine ganze algebraische Zahl, also $\frac{n a}{r}$ eine ganze algebraische Zahl; nun ist $\frac{n}{r}$ keine ganze Zahl, daher enthält r einen Primfaktor q, der in der kritischen Zahl a aufgeht. Also ist nach Satz 29 $a = -2$, $r = 2$, womit Satz 32 bewiesen ist.

Durch Satz 31 und Satz 32 erhalten wir zusammenfassend

Satz 34. *Für das Polynom $g(x) = x^r + a$, $r \geq 2$, $a \neq 0$ ist die Folge V_g der Iterierten von g im allgemeinen vollständig, d.h. außer den Iterierten von g gibt es kein Polynom $f(x)$ positiven Grades, das mit $g(x)$ vertauschbar ist. Die einzige Ausnahme ist $g(x) = x^2 - 2$, das das quadratische Polynom der Grundkette* (II) *in § 7 ist. V_g ist nur eine Teilfolge der Kette* (II).

Der Fall $a=0$, $g(x)=x^r$, $r\geq 2$ erledigt sich leicht. V_g ist nicht vollständig. Denn alle Lösungen von

$$fg=gf, \quad f(x^r)=(f(x))^r$$

sind nach den Bemerkungen, die im Anschluß an Satz 19 gemacht wurden:

$$f(x)=Cx^m, \quad C^{r-1}=1, \quad m=0,1,2,\ldots,$$

wozu noch $f(x)\equiv 0$ kommt.

§ 11. Quadratische Polynome.

Ist $g(x)$ ein beliebiges Polynom positiven Grades und V_g die zugehörige Folge (4) der Iterierten von g, und transformiert man durch eine Ähnlichkeitstransformation T, (§ 5), g in g_1, so geht bei Anwendung desselben T auf alle Elemente von V_g diese Folge in V_{g_1} über. Die Folgen heißen ähnlich. Ist die eine vollständig, so ist die andere auch vollständig. Nun sei

$$g(x)=a_0x^2+a_1x+a_2, \qquad a_0\neq 0, \tag{61}$$

ein gegebenes quadratisches Polynom. Bezeichnet man die Diskriminante von $g(x)$ mit D, also

$$D=a_1^2-4a_0a_2, \tag{61$_1$}$$

und setzt man

$$I(g)=I=D-2a_1=a_1^2-4a_0a_2-2a_1, \tag{61$_2$}$$

so ist I eine Invariante der Transformation. Geht also g durch die Transformation $T(c,d)$, § 5, Gl. (22), in

$$g_1(x)=b_0x^2+b_1x+b_2 \tag{62}$$

über und ist

$$\Delta=b_1^2-4b_0b_2, \tag{62$_1$}$$

$$I(g_1)=\Delta-2b_1=b_1^2-4b_0b_2-2b_1, \tag{62$_2$}$$

so ist

$$I(g)=I(g_1) \tag{63}$$

und umgekehrt folgt aus (63), daß (61) und (62) ähnliche Polynome sind, die durch die Transformation $T(c,d)$ mit

$$c=\frac{b_0}{a_0}, \qquad d=\frac{b_1-a_1}{2a_0} \tag{64}$$

ineinander transformiert werden, d.h. es ist

$$g_1=T^{-1}fT(x). \tag{65}$$

Nun gibt es nach § 5 eine einzige Transformation T, die $g(x)$ in $g_1(x)=x^2+a$ transformiert. Es ist also bei Anwendung dieser Transformation

$$b_0=1, \quad b_1=0, \quad b_2=a, \quad \Delta=-4a, \quad I(g_1)=-4a.$$

Also

$$a_1^2 - 4a_0a_2 - 2a_1 + 4a = 0. \tag{63_1}$$

Wegen Satz 34 und der anschließenden Bemerkungen ist V_{g_1} eine vollständige Folge dann und nur dann, wenn $a \neq 0$, $a \neq -2$; also ist V_g dann und nur dann vollständig, wenn

$$I(g) = a_1^2 - 4a_0a_2 - 2a_1 \neq 8 \quad \text{und} \quad \neq 0$$

ist. Ist dagegen $I(g) = 0$ oder 8, so liegt g in einer V-Kette, die zu (I) oder (II), s. § 7, ähnlich ist. Wir bekommen so

Satz 35. *Ist $g(x) = a_0x^2 + a_1x + a_2$, $a_0 \neq 0$, ein quadratisches Polynom, so sind im allgemeinen die einzigen Polynome positiven Grades, die mit $g(x)$ vertauschbar sind, die Iterierten von $g(x)$. Nur wenn $I(g) = a_1^2 - 4a_0a_2 - 2a_1 = 0$ oder $= 8$ ist, gibt es für jeden positiven Grad m genau ein Polynom $f(x)$, das mit $g(x)$ vertauschbar ist. Man bekommt es im Falle $I(g) = 0$ aus dem Gliede x^m der Kette* (I) *und im Falle $I(g) = 8$ aus dem Gliede $f_m(x)$ der Kette* (II) *durch eine Ähnlichkeitstransformation.*

Für die quadratischen Polynome $g(x)$ und die reinen Polynome ist in den beiden letzten Paragraphen die Frage nach der Vollständigkeit der Folge der Iterierten von $g(x)$ entschieden worden. Man kann sehr leicht Polynomklassen angeben, bei denen die zugehörige Folge V_g nicht vollständig ist. Ist nämlich ε eine r-te Einheitswurzel $\neq 1$ und

$$g(x) = \sum_{\lambda=0}^{q} a_{\lambda r+1} x^{\lambda r+1}$$

ein Polynom des Grades $m = qr + 1$, so ist

$$g(\varepsilon g(x)) = \varepsilon \sum_{\lambda=0}^{q} a_{\lambda r+1} (g(x))^{\lambda r+1} = \varepsilon g(g(x)).$$

Es ist also

$$f(x) = \varepsilon g(x)$$

mit $g(x)$ und daher mit jedem Element von V_g vertauschbar, ohne daß $f(x)$ ein Element von V_g ist. V_g ist also nicht vollständig.

Matematisk Institutt, Norges Tekniske Høgskole, Trondheim (Norwegen).

(Eingegangen am 18. März 1955.)

KELLER, O.-H.
Math. Zeitschr. Bd. 63, S. 277—285 (1955)

Eine Bemerkung zur Ausführung der körpertheoretischen Operationen in erträglich vielen Schritten.

Dem Gedächtnis von ISSAI SCHUR.

Von

OTT-HEINRICH KELLER.

Wohl die eindrucksvollste Vorlesung, die ich als Student hörte, war die von I. Schur im Wintersemester 1926/27 über Galoissche Theorie. Es sei mir daher gestattet, meine Verehrung für ihn durch eine bescheidene Bemerkung zur Galoisschen Theorie zu bezeugen.

Wie allen Studenten war es auch mir eine große Enttäuschung, daß die Galoissche Theorie auch im Fall einer auflösbaren Gruppe die Berechnung der Wurzeln zwar theoretisch, aber nicht praktisch möglich macht. Und mir ist das Bild und der Tonfall des Augenblicks noch gegenwärtig, als Schur sagte, daß die Galoissche Resolvente einer Gleichung 5. Grades im allgemeinen vom 60. Grade sei. Es ist nicht der hohe Grad allein, der die Rechnung so sehr verwickelt. Sondern es werden, etwa beim Satz vom primitiven Element, Unbestimmte eingeführt, und man muß dann mit Polynomen in mehreren Veränderlichen rechnen. Es wird zwar gezeigt, daß es immer Zahlen gibt, durch die man diese Unbestimmten ersetzen darf, aber das sind reine Existenzaussagen und diese werden in keinem der üblichen Lehrbücher zu einem Konstruktionsverfahren ausgebaut. Hier einen noch so kleinen und unscheinbaren Schritt weitergehen zu können, scheint mir eine Mitteilung zu rechtfertigen. Der Grundgedanke findet sich in einer Arbeit von I. SCHUR[1]). Die angegebene Vereinfachung wird uns im IV. Abschnitt in den Stand setzen, die Galoissche Gruppe einer numerisch gegebenen Gleichung 5. Grades mit erträglichem Aufwand zu berechnen.

I.

Dieser kleine Schritt ist in dem folgenden Satz enthalten.

Satz. Es sei G ein Integritätsbereich, K sein Quotientenkörper. K sei separabel und enthalte unendlich viele Elemente. Die Zahlen α und β seien ganz algebraisch über G; β möge der irreduziblen Gleichung $f(\beta)=0$ über G genügen. *Es sei u eine Zahl aus G, die in der Diskriminante von f nicht aufgeht. Dann ist $\beta+u\alpha$ ein primitives Element des Körpers $K(\alpha,\beta)$.*

Um den Satz zu beweisen, denken wir uns die Konjugierten $\beta=\beta_1, \beta_2, \ldots, \beta_r$ und $\alpha=\alpha_1, \alpha_2, \ldots, \alpha_s$ von β und α hingeschrieben. Angenommen,

[1]) SCHUR, I.: Zur Irreduzibilität der Kreisteilungsgleichung. Math. Z. 29, 463 (1928).

unter den Größen $\beta_i + u\alpha_k$ $(i = 1, \ldots, r;\ k = 1, \ldots, s)$ seien zwei gleiche, $\beta_{i_1} + u\alpha_{k_1} = \beta_{i_2} + u\alpha_{k_2}$, so wäre $\beta_{i_1} - \beta_{i_2} = u(\alpha_{k_2} - \alpha_{k_1})$, und u wäre ein Teiler von $\beta_{i_1} - \beta_{i_2}$ und damit auch von der Diskriminante von f. Also sind die $\beta_i + u\alpha_k$ alle verschieden. Daraus kann man in der üblichen Weise schließen, daß $\beta + u\alpha$ ein primitives Element des Körpers $K(\alpha, \beta)$ ist.

II.

Es sei nun ein Polynom $g(x; \alpha)$ mit Koeffizienten aus $G[\alpha]$ gegeben. Will man $g(x; \alpha)$ in irreduzible Faktoren zerlegen, so bildet man üblicherweise die Norm $F(x; u) = N(g(x - u\alpha; \alpha))$ mit einer Unbestimmten u und zerlegt F in Faktoren $F = F_1 \cdot F_2 \ldots F_n$. Dann sind die Faktoren von g die größten gemeinsamen Teiler der F_i mit g[2]). Mit unbestimmten u ist diese Rechnung auch in einfachen Fällen nicht mehr in erträglich vielen Schritten zu bewältigen. Ist α etwa eine Irrationalität 5. Grades, so bedeutet die Normbildung die Berechnung einer Determinante 5. Grades, deren Elemente Polynome in x und u sind. Weiter hätten wir dann die Zerlegung eines Polynoms sehr hohen Grades in zwei Veränderlichen in Faktoren zu zerlegen. Für die Rechnung ist es daher sehr vorteilhaft zu wissen, *daß das u dabei durch irgendeine Zahl ersetzt werden darf, die nur nicht in der Diskriminante eines irreduziblen Teiles von $N(g(x; \alpha))$ aufgeht.* Zum Beweise bezeichnen wir die Konjugierten von α mit $\alpha_1, \ldots, \alpha_s$, die Nullstellen von $g(x, \alpha_k)$ mit $\beta_i^{(k)}$ $(i = 1, \ldots, r;\ k = 1, \ldots, s)$. Dann ist $g(x, \alpha_k) = \prod_{i=1}^{r} (x - \beta_i^{(k)})$ und $N(g(x; \alpha_k)) = \prod_{i=1}^{r} \prod_{k=1}^{s} (x - \beta_i^{(k)})$, also die Gleichung für $\beta_i^{(k)}$ über K oder ein Vielfaches von ihr. Es ist $F(x; u) = \prod_{i=1}^{r} \prod_{k=1}^{s} (x - u\alpha_k - \beta_i^{(k)})$, zerfällt also nach dem bewiesenen Satz in die irreduziblen Polynome, deren Nullstellen die primitiven Elemente der Körper $K_{ik}(\alpha_k, \beta_i^{(k)})$ sind. Der größte gemeinsame Teiler eines solchen Teilpolynoms mit $g(x - u\alpha, \alpha)$ hat als Nullstellen genau eine solche Reihe von Größen $u\alpha + \beta_k^{(1)}$ $(k = k_1, k_2, \ldots)$ die über $K(\alpha)$ konjugiert sind, ist also irreduzibel.

Ist die Diskriminante von $N(g(x; \alpha))$ nicht ohne weiteres berechenbar, so bilde man $N\left(u^r g\left(\frac{x - \alpha}{u}; \alpha\right)\right) = \prod_{k=1}^{s} \prod_{i=1}^{r} (x - \alpha_k - u\beta_i^{(k)})$, wo u jetzt in der Diskriminante der irreduziblen Gleichung $f(\alpha) = 0$ für α nicht aufgehen möge.

III.

Es sei jetzt eine irreduzible Gleichung $f(x) = 0$ vom n-ten Grade vorgelegt, die wir mit Hilfe der Galoisschen Theorie behandeln sollen; es sei α eine Wurzel. Wir können die Gleichung $g(x; \alpha) = \frac{f(x)}{x - \alpha}$ für die übrigen Wurzeln bestimmen. Es sei β eine Wurzel von g und damit auch von f. Eine Gleichung für ein primitives Element des Körpers $K(\alpha, \beta)$ ist dann ein Teiler von $F(x) = N(g(x - u\alpha; \alpha))$, wo u nur in der Diskriminante von f nicht aufgehen möge.

[2]) WAERDEN, VAN DER: Moderne Algebra I, 1. Aufl., § 37, S. 128. 1930.

Diese Norm ist in den einfachsten nicht trivialen Fällen noch mit erträglichem Aufwand zu berechnen. Dieses F ist freilich noch nicht die Galoissche Resolvente, sondern erst eine Wiederholung des Verfahrens führt zu ihr hin. Es scheint aber zweckmäßig, erst einmal das F auszunutzen. Wir erhalten wesentliche Aufschlüsse über die Galoissche Gruppe, wenn wir F in irreduzible Faktoren zerlegen. Diese Faktorzerlegung mag mühevoll sein; sie ist aber nach einem Satz von KRONECKER[3]) prinzipiell möglich. Die dort angegebene Methode ist aber praktisch nicht ausführbar, und wir werden uns für unseren Fall andere Ansätze suchen müssen. Zunächst ist F im Restklassenkörper mod u zerlegbar in $(f(x))^{n-1}$, denn es ist $N(g(x;\alpha)) = (f(x))^{n-1}$. Die Teiler von F werden also als Grad ein Vielfaches von n haben. Suchen wir eine Zerlegung $F = P\,Q$ in Polynome bestimmter, durch n teilbarer Grade, so denken wir uns, wenn u eine Primzahl ist, P und Q u-adisch entwickelt: $P = \sum_i p_i u^i$, $Q = \sum_i q_i u^i$. Dann sind uns p_0 und q_0 als Potenzen von f bekannt. Wir können nun die Koeffizienten der verschiedenen Potenzen von u rechts und links miteinander vergleichen. Wir bilden zunächst $\frac{1}{u}(F - p_0 q_0) \equiv p_1 q_0 + q_1 p_0 \pmod u$ und können p_1 und q_1 aus dieser Kongruenz, wenn überhaupt, dann linear berechnen. Die Vieldeutigkeit in der Bestimmung von p_1 und q_1 können wir ausnutzen, um die nächste Kongruenz

$$\frac{1}{u}\left(\frac{1}{u}(F - p_0 q_0) - p_1 q_0 - p_0 q_1\right) - p_1 q_1 \equiv p_2 q_0 + q_2 p_0 \pmod u$$

lösbar zu machen, wenn dies überhaupt möglich ist. Diese Kongruenz ist doch dann und nur dann lösbar, wenn die linke Seite durch den größten gemeinsamen Teiler von p_0 und q_0 teilbar ist. So fortfahrend kommen wir entweder auf eine unlösbare Kongruenz oder auf eine Zerlegung von F.

Ist über dem Grundkörper keine solche Zerlegung möglich, so ist die Gruppe zweifach transitiv. Man wird nun eine Zerlegung im Körper $K(\sqrt{D})$ ansetzen, wo D die Diskriminante ist. Dabei ist die angegebene Methode etwas zu modifizieren: Die Faktoren von F müssen in $K(\sqrt{D})$ konjugiert sein, also die Form $P + Q\sqrt{D}$ und $P - Q\sqrt{D}$ haben. Ist nun d irgendein Primteiler von D so können wir auch mod d rechnen. Wir nehmen an, d sei der einzige Primfaktor, der in einer ungeraden Potenz in D aufgeht. Es ist $F \equiv p_0^2 \pmod d$. p_0 ist also, wenn überhaupt, dann leicht zu berechnen.

Bei beiden Moduln u und d besteht die Hauptschwierigkeit in der Berechnung des Anfangsgliedes der Entwicklung von Q. Dazu ist eine quadratische Kongruenz zu lösen. Man wird hier im allgemeinen nicht durchkommen, ohne eine gewisse Anzahl von Fällen durchzuprobieren. Wir arbeiten z.B. mit dem Modul d. Angenommen, wird könnten die ersten p_i und q_i berechnen, und es sei $q_k u^k$ das erste Glied in der Entwicklung von Q, das nicht identisch verschwindet. Wir setzen $\frac{1}{d}(F - p_0^2) = F_1$. Angenommen wir hätten $p_0, p_1,$

[3]) Grundzüge einer arithmetischen Theorie der algebraischen Größen. § 4 [Crelles J. 92 (1882)].

$\ldots, p_{i-1}$ $(i \leq 2h+1)$ sowie $F_1, \ldots, F_{i-1}$ schon bestimmt, und es sei $F_{i-1} \equiv \sum_{\nu=0}^{i-1} p_\nu p_{i-1-\nu} \pmod d$. Wir setzen $\frac{1}{d}\left(F_{i-1} - \sum_{\nu=0}^{i-1} p_\nu p_{i-1-\nu}\right) = F_i$. Die Betrachtung der nächsten Glieder ergibt $F_i \equiv 2p_0 p_i + \sum_{\nu=1}^{i-1} p_\nu p_{i-\nu} \pmod d$. Wir setzen $G_i = F_i - \sum_{\nu=1}^{i-1} p_\nu p_{i-\nu}$ und errechnen p_i aus $2p_0 p_i \equiv G_i \pmod d$. So fortfahrend kommen wir zu der Kongruenz $F_{2h+1} = 2p_0 p_{2h+1} + \sum_{\nu=1}^{2h} p_\nu p_{2h+1-\nu} - q_h^2 \pmod d$, also $q_h^2 \equiv -G_{2h+1} \pmod{p_0, d}$. Nun ist zwar mit Hilfe des quadratischen Reziprozitätsgesetzes[4]) leicht zu entscheiden, ob diese Kongruenz lösbar ist oder nicht. Die wirkliche Lösung kann aber erhebliche Schwierigkeiten machen. Ist p_0 ein Primpolynom vom Grade k, so können wir etwa das Galoisfeld heranziehen, dessen erzeugendes Element γ durch $p_0(\gamma) = 0 \pmod d$ definiert ist. Es seien $\gamma = \gamma_1, \gamma_2, \ldots, \gamma_k$ die verschiedenen Wurzeln von $p_0 = 0 \pmod d$. Das primitive Element der multiplikativen Gruppe dieses Galoisfeldes sei δ. Dann können wir die Kongruenzen $q_h^2(\gamma_i) \equiv -G_{2h+1}(\gamma_i)$ aufstellen. Nun ist die Zahl $-G_{2h+1}(\gamma_i) = \delta^N$, da die multiplikative Gruppe immer zyklisch ist. Ist dabei d und N eine ungerade Zahl, so ist die Kongruenz unlösbar; ist N gerade, so ist $q_h(\gamma_i) = \delta^{\frac{1}{2}N}$ zu setzen. Wenn wir die Werte von q_h an den k Stellen γ_i kennen, können wir mit Hilfe der Lagrangeschen Interpolationsformel $q_h(x)$ berechnen.

Bei den riesigen Galoisfeldern, die wir zu erwarten haben, wird aber das Auffinden der Zahl N im allgemeinen eine aussichtslose Aufgabe sein: Ist $f(x)$ vom 5. Grad, so ist $F(x)$ vom 20. Grad, p_0 vom 10. Grad und das Galoisfeld hat d^{10} Elemente. N müssen wir also durch Probieren aus den Zahlen $1, \ldots, d^{10}$ aussuchen. Die Aufgabe vereinfacht sich wesentlich, wenn p_0 kein Primpolynom ist.

Ist q_h bekannt, so setzen wir für $i > 2h+1$

$$\frac{1}{u}\left(G_{i-1} - 2p_0 p_{i-1} + 2q_h q_{i-h-2}\right) = F_i$$

$$F_i - \sum_{\nu=1}^{i-1} p_\nu p_{i-\nu} + \sum_{\nu=h+1}^{i-h-1} q_\nu q_{i-\nu-1} = G_i$$

und können, wenn p_0 und q_0 teilerfremd sind, q_{i-h} und p_i aus $2p_0 p_i - 2q_h q_{i-h} \equiv G_i \pmod d$ linear berechnen.

Unter der i-ten Stufe wollen wir dabei den Vergleich der Koeffizienten von d^i verstehen.

IV.

Als Anwendung stellen wir uns die Aufgabe, die Gleichung $x^5 + 5x^3 + 5x - 1 = 0$ zu behandeln, ihre Gruppe zu bestimmen, und wenn sich diese als auflösbar herausstellen sollte, die Gleichung aufzulösen.

a) Die Gleichung ist irreduzibel, wie man nach der Substitution $x = y + 1$ nach Eisenstein leicht sieht.

b) Die Diskriminante der Gleichung ist 5^7, wie eine kurze Rechnung zeigt.

[4]) HASSE: Zahlentheorie, § 5g, S. 72. 1949.

c) Die Gleichung gestattet in Restklassenkörpern bzw. in den zugehörigen durch die Gleichung $x^2+x-1\equiv 0$ definierten Galoisfeldern folgende Zerlegungen in Primfaktoren; wenn ϱ und $\bar{\varrho}$ die Wurzeln von $x^2+x-1\equiv 0$ sind:

$$\begin{aligned} f(x) &\equiv (x+1)(x^4+x^3+1) \pmod 2 \\ &\equiv (x+1)(x^2+\varrho x+\varrho)(x^2+\bar{\varrho}x+\bar{\varrho}) \\ f(x) &\equiv (x+1)(x^4-x^3-1) \pmod 3 \\ &\equiv (x+1)(x^2+\varrho x+1+\varrho)(x^2+\bar{\varrho}x+1+\bar{\varrho}) \\ f(x) &\equiv (x-1)^5 \pmod 5 \\ f(x) &\equiv (x-3)(x^4+3x^3+5) \pmod 7 \\ &\equiv (x-3)(x^2-3\varrho x+5+\varrho)(x^2-3\bar{\varrho}x+5+\bar{\varrho}) \end{aligned}$$

$f(x)$ ist irreduzibel (mod 11).

Die Gruppe enthält also sicher einen Fünfer- und einen Viererzyklus, ihr alternierender Normalteiler ein Paar von Zweierzyklen[5]). Es kommen also nur die volle symmetrische und die volle metazyklische Gruppe in Betracht. Eine Entscheidung zwischen ihnen können wir durch Betrachtung von Kongruenzen nicht erreichen.

d) Es ist $g(x;\alpha)=x^4+\alpha x^3+(\alpha^2+5)x^2+(\alpha^3+5\alpha)x+\alpha^4+5\alpha^2+5$. Die kleinste Zahl, die nicht in der Diskriminante aufgeht, ist 2; wir berechnen

$$g(x-2\alpha;\alpha)=11\alpha^4-23\alpha^3x+\alpha^2(19x^2+15)+\alpha(-7x^3-15)x+x^4-5x^2-5.$$

Die Norm $N\big(g(x-2\alpha;\alpha)\big)$ finden wir, indem wir $g(x-2\alpha;\alpha)$ nacheinander mit $1, \alpha, \alpha^2, \alpha^3, \alpha^4$ multiplizieren, mod $f(\alpha)$ auf Polynome 4. Grades reduzieren und die Determinante der Koeffizienten der verschiedenen Potenzen von α bilden. Es wird $N\big(g(x-2\alpha;\alpha)\big)=$

11	$-23x$	$19x^2+15$	$-7x^3-15x$	x^4+5x^2+5
$-23x$	$19x^2-40$	$-7x^3-15x$	x^4+5x^2-50	11
$19x^2-40$	$-7x^3+100x$	x^4+5x^2-50	$115x+11$	$-23x$
$-7x^3+100x$	x^4-90x^2+150	$115x+11$	$-95x^2-23x+200$	$19x^2-40$
x^4-90x^2+150	$35x^3-385x+11$	$-95x^2-23x+200$	$35x^3+19x^2-500x-40$	$-7x^3+100x$

Bei der Berechnung dieser Determinante kann man mit Hilfe des Laplaceschen Satzes nach den Elementen der beiden ersten und der drei letzten Spalten entwickeln. Andere Determinantensätze heranzuziehen scheint sich nicht zu lohnen. Die Rechnung ist im Prinzip einfach, aber freilich etwas umfangreich. Wir erhalten

$$\begin{aligned} N\big(g(x-2\alpha;\alpha)\big) &= F(x) \\ &= x^{20}+80x^{80}+2570x^{16}+78x^{15}+43200x^{14}+3930x^{13}+417275x^{12}+ \\ &\quad +72420x^{11}+2418344x^{10}+652050x^9+8532510x^8+3080100x^7+ \\ &\quad +18139840x^6+7500972x^5+22191925x^4+8767440x^3+ \\ &\quad +13438600x^2+4341810x+2328551. \end{aligned}$$

[5]) Waerden, van der: Moderne Algebra I, 1. Aufl., § 56, S. 188. 1930.

e) Die Kongruenzen unter c) zeigen, daß mit einer Zerlegbarkeit im rationalen Zahlkörper nicht zu rechnen ist, wohl aber im Körper $K(\sqrt{5})$ und dann notwendig in Polynome gleichen Grades der Form $F=(P'+\varrho Q')\,(P'+\bar{\varrho}Q')$, wo ϱ und $\bar{\varrho}$ die Wurzeln von $x^2+x-1=0$, nämlich $\varrho, \bar{\varrho}=\frac{-1\pm\sqrt{5}}{2}$ sind. Weil $F(x)$ mit $1\cdot x^{20}$ beginnt, dürfen wir annehmen, daß beide Faktoren mit $1\cdot x^{10}$ beginnen, daß also Q keine Glieder 10. Ordnung enthält. Das Kroneckersche Verfahren ist hier überhaupt nicht unmittelbar anzuwenden, weil der Körper unendlich viele Einheiten, jede Zahl also unendlich viele Faktoren hat.

Nach III. haben wir mod 2 die Zerlegung $F(x)=(x^5+x^3+x+1)^4$ zu erwarten. Sie ist in der Tat eine mögliche Zerlegung. Also sind alle Koeffizienten von Q' gerade und wir können in Zukunft $\sqrt{5}$ an Stelle von ϱ als Glied der Körperbasis wählen. Damit rechnet sich einfacher. Wir suchen also eine Zerlegung $F=(P+Q\sqrt{5})\,(P-Q\sqrt{5})$, wo $P=P'-\frac{Q'}{2}$, $Q=\frac{Q'}{2}$. Die weitere Rechnung mod 2 hat zwei Nachteile: Erstens geht sie sehr langsam und zweitens wird das Verfahren mod 8 nicht mehr linear: Wir erhalten

$$F\equiv p_0^2+4p_1p_0+4p_1^2-q_0^2-4q_1q_0-q_1^2 \pmod 8,$$

und dies ist eine Kongruenz 2. Grades für p_1 und q_1. Die Rechnung wurde deshalb 5-adisch weitergeführt. Wir berechnen p_0 aus der Kongruenz $p_0^2=F \pmod 5$ zu $p_0=(x^5+2)^2\equiv(x+2)^{10} \pmod 5$. Das ist außerordentlich günstig für die Lösung der quadratischen Kongruenz für das Anfangsglied der Entwicklung von Q. Da wir häufig durch p_0 zu dividieren haben, läge es nahe, $F(x)$ durch $x=y-2$ in die Form $F(y)$ zu transformieren. Jedoch ist die Umrechnung eines Polynoms 20. Grades rechnerisch nicht in wenigen Augenblicken durchzuführen und führt zu Koeffizienten der Größenordnung 10^{11} und 10^{12}, mit denen man auf den üblichen Rechenmaschinen nicht mehr alle Operationen ohne weiteres ausführen kann. Wir verzichten daher darauf und transformieren dafür lieber jedes $G_i(x)$ für sich mod 5 in $G_i^*(y)$. Diese Rechnung ist jedesmal in wenigen Zeilen durchzuführen. Wir geben nun eine Tabelle der $G_i^*(y)$. Wir müssen dabei für jedes i eine bestimmte vorläufige Lösung der p_ν und q_ν zugrunde legen. Im weiteren Verlauf der Rechnung werden wir noch zu Korrekturen gezwungen und legen demzufolge andere vorläufige Lösungen zugrunde. Es ist also

	y^{19}	y^{15}	y^{10}	y^5	y^0
G_1^*	. 1 4 1 3	. . . . .	. . . . .	. . . . .	
G_2^*	. 2 3 2 .	1 . 3 1 2	. . . . .	. . . . .	
G_3^*	. . . 1 3	4 4 3 1 .	. . . . .	. . . . .	
G_4^*	. . . . .	1 2 2 3 4	. . . . .	. . . . .	
G_5^*	. . . . .	 3	. . . . .	. . . . .	
G_6^*	. . . . .	 1	. 2 3 . .	. . . . .	
G_7^*	. . . . .	. . . 2 4	1 2 2 4 1	. . .	
G_8^*	. . . . .	 1	1 3 4 4 .	4 2 . .	
G_9^*	. . . . .	 1	3 3 3 1 1	. . . . .	
G_{10}^* bis G_{19}^*	. . .	. . . . .	. . . . .	. . . . .	

Aus den Kongruenzen ergeben sich daraus für p_i und q_i

	x^{10}	x^5				x^0	
p_0	1	 4	.	.	.	. 4	
p_1	.	. 3 . 2 2	.	1	.	4 4	
p_2	.	. 1 . 4 1	3	1	1	3 3	
p_3	.	. . . 3 .	2	3	4	3 .	$+2r_1^2$
p_4	.	. . . . .	3	.	4	1 .	$+4r_1r_2$
p_5	.	. . . . .	.	.	.	. 4	$+4r_1r_3+2r_2^2$
p_6	.	. . . . .	.	.	.	. 3	$+4r_2r_3+2r_1r_4$
p_7	.	. . . . .	.	.	.	−4 2	$+2r_2r_4+2r_3^2+2(x+1)r_5+(-4x+5)s$
p_8	.	. . . . .	.	.	.	. −2	$+4xr_4+3r_5+2(x+1)r_6$
p_9	.	. . . . .	.	.	.	. 1	$+2(x+1)r_7$
q_1			.	.	.	. .	$(x^5+2)\,2r_1$
q_2			.	.	.	. .	$A_1r_1+2(x^5+2)r_2$, wo $A_1=3x^3+2x+2$
q_3			.	.	.	. .	$A_2r_1+A_1r_2+2(x^5+2)r_3$, wo $A_2=x^3+3x+1$
q_4			1	5	13	19 7	$+A_1r_3+(x^5+2)r_4$
q_5			.	.	1	−1 1	$+A_2r_3-2A_1r_4+(x+2)^3r_5+(x^3+7x^2+10x+7)s$
q_6			.	.	1	4 4	$+(4x^3+2x^2+4x)r_4+(2x+4)r_5+(x+2)^3r_6$
q_7			.	.	2	4 3	$+(x+2)^3r_7$

Bei der Suche nach der allgemeinen Lösung der 1-ten Stufe müssen wir die Kongruenz $q_0^2 = 0 \pmod{5, p_0}$ lösen. Sie ergibt $q_0 = (x^5+2)\,2r_0$. Zur Bestimmung von r_0 erweist sich eine Fallunterscheidung als notwendig nach der Potenz von y, durch die dieses r_0 als teilbar angenommen wird. Die Potenzen y^0 und y^1 erweisen sich bereits auf der zweiten Stufe als unmöglich, y^2 als auf der dritten, y^3 und y^4 auf der vierten. Wir müssen also mit $r_0 = 0$ weiterrechnen. Auf der dritten Stufe stehen wir wieder vor der Kongruenz $q_1^2 = 0 \pmod{5, p_0}$ mit der allgemeinen Lösung $q_1 = (x^5+2)\,2r_1$. Auf jeder folgenden Stufe wurde dann eine Partikularlösung berechnet und jedesmal die allgemeine Lösung der homogenen Kongruenz

$$2(x^5+2)^2 p_{i+1} - 4(x^5+2)\,r_1 q_i \equiv 0 \pmod 5$$

mit unbestimmten r_i hinzugefügt. Bei den ersten Stufen wurde zunächst einmal versuchsweise auf die volle Allgemeinheit verzichtet und außer acht gelassen, daß r_1 und p_0 noch einen Faktor gemein haben können. Erst nachdem sich auf der 6. Stufe herausgestellt hat, daß auch r_1 durch y^2 teilbar ist, wurde dies berücksichtigt. Zu den G_i treten dann noch gewisse Ausdrücke, die sich aus den r_i zusammensetzen. Diese unbestimmten Polynome r_i lassen sich dann auf den höheren Stufen aus der Bedingung berechnen, daß G_i durch den größten gemeinsamen Teiler von $r_1(x^5+2)$ und p_0 teilbar sein muß, wenn die Kongruenzen überhaupt noch Lösungen haben sollen.

Auf der 6. Stufe ergab sich die Kongruenz $y^{10} + 2y^8 + 3y^7 \equiv G_6 = -4q_1q_4 + 2p_0p_6 + r_1^2y^4 + 4r_1^4 + 4A_1r_1r_3y^5 + 4A_2r_1r_2y^5 - 8r_2r_3y^{10} \pmod 5$. Daraus können wir schließen, daß r_1 genau durch y^2 teilbar ist. Im übrigen ist die Kongruenz von r_1 im 4. Grad, und wir werden sie nicht in allgemeinen Buchstaben lösen können, sondern wir werden für r_1 verschiedene Ansätze durchprobieren

müssen. Am nächsten liegt wohl der, daß die Glieder unter y^{10} für sich durch r_1 teilbar sind, daß also $r_1 = \lambda(2y+3)y^2$. Auf der 7. Stufe findet man dann $\lambda^2 = 4$. Wir können uns willkürlich für eine der beiden Wurzeln, etwa für $\lambda = 3$ entscheiden; $\lambda = 2$ liefert den anderen Faktor der Zerlegung von F. Wir setzen dementsprechend $r_1 = (x+1)(x+2)^2$. Die weitere Rechnung verläuft nun linear. Auf der 7. Stufe finden wir $r_2 = 4x + 1 + sy^2$; die 8. Stufe liefert $s = 4$. Unter den Repräsentanten der entsprechenden Restklassen wählen wir $r_2 = -x^2 - 3$. Auf der 8. Stufe findet man noch $r_3 = -2 - r_5$, auf der 9. Stufe $r_3 = -2$, $r_4 = 1$, $r_5 = 0$, $r_6 = -2$. Die Repräsentanten der Restklassen sind dabei so gewählt, daß jedesmal die höchsten Glieder von $G_i(x)$ auf beiden Seiten gleich und nicht nur (mod 5) kongruent werden.

Beim Weiterrechnen werden alle $G_i = 0$ $(i \geq 10)$ und F_{19} verschwindet, wenn wir alle r_i $(i \geq 7)$ durch Null ersetzen. Wir haben also eine Zerlegung $F = (P + Q\sqrt{5})(P - Q\sqrt{5})$, wobei

$$P = x^{10} + 40x^8 + 735x^6 + 39x^5 + 6200x^4 + 2555x^3 + 19025x^2 + \\ + 7095x + 20974$$

$$Q = 10x^8 + 280x^7 + 15x^5 + 2650x^4 + 495x^3 + 8400x^2 + 3135x + 9355.$$

f) Wir haben nun den größten gemeinsamen Teiler von $(P + Q\sqrt{5})$ mit $g(x - 2\alpha; \alpha)$ zu suchen. Wir werden dabei $\sqrt{5}$ wieder durch $2\varrho + 1$ mit $\varrho = \frac{-1+\sqrt{5}}{2}$ ersetzen müssen. Mit Hilfe des Euklidischen Algorithmus finden wir nach zwei Divisionen für diesen größten gemeinsamen Teiler das Polynom

$$x^2 - (4 + \varrho)\alpha x + \alpha^2(5 + 2\varrho) + 3 + \varrho,$$

also, wenn wir $x - 2\alpha$ wieder durch x ersetzen

$$g(x) = (x^2 - \varrho\alpha x + \alpha^2 + 3 + \varrho)(x^2 - \bar{\varrho}\alpha x + \alpha^2 + 3 + \bar{\varrho}) = g_1 g_2\ ^{6}).$$

g) Die Zerlegung aus f) zeigt uns, daß die Gruppe von $f(x)$ nicht die symmetrische, sondern die volle metazyklische ist. Es sei β eine Wurzel von g_1; dann ist $\varrho\alpha - \beta$ die andere Wurzel. Da es nun keine Permutation der Wurzeln von $f(x)$ geben kann, die α und die Wurzeln von g_1 festläßt, die von g_2 aber vertauscht, muß g_2 bei Adjunktion von β zerfallen. Um die Faktorzerlegung zu finden, überzeugen wir uns zunächst davon, daß es ein Element der Gruppe gibt, das α und β vertauscht. In der Tat gibt es Gruppenelemente mit Zweierzyklen und wegen der Transitivität auch ein solches, das α mit einer anderen Wurzel γ vertauscht. Ist γ eine Wurzel von g_1 so sind wir fertig. Ist γ eine Wurzel von g_2, so beachten wir, daß es einen Automorphismus des Körpers

[6]) Dieses Ergebnis hätten wir auch bekommen, wenn wir für $g(x; \alpha)$ eine Zerlegung in den beiden Unbestimmten x und α angesetzt hätten. In unserem speziellen Beispiel hat sich also der Weg über die Norm als Umweg erwiesen.

gibt, bei dem $\sqrt{5} \to -\sqrt{5}$, $g_1 \to g_2$, $\alpha \to \alpha$, $\gamma \to \beta$. Also gibt es ein Gruppenelement, das den Zweierzyklus $(\alpha\beta)$ enthält. Dieses Element führt $\varrho\alpha - \beta$ in $\varrho\beta - \alpha$ über.

$\varrho\beta - \alpha$ ist von α, β und $\varrho\alpha - \beta$ verschieden, weil sich sonst eine lineare Beziehung zwischen α und β ergeben würde, während g_1 doch irreduzibel ist. Also ist $\varrho\beta - \alpha$ die eine Wurzel von g_2; $-\varrho(\alpha+\beta)$ ist dann die andere. Die Wurzeln 1. α, 2. β, 3. $\varrho\alpha - \beta$, 4. $-\varrho(\alpha+\beta)$, 5. $\varrho\beta - \alpha$ gestatten offensichtlich die Permutationen (12) (35) und (23) (45). Diese erzeugen die alternierende Untergruppe der metazyklischen Gruppe. Jetzt liefern die bekannten Methoden[7]) die Wurzeln $\alpha_\nu = \varepsilon^\nu \sqrt[5]{\frac{1+\sqrt{5}}{2}} + \varepsilon^{5-\nu} \sqrt[5]{\frac{1-\sqrt{5}}{2}}$ ($\nu = 0, 1, 2, 3, 4$ und $\varepsilon^2 = 1$).

[7]) PERRON: Algebra II, 2. Aufl., § 43, S. 226. 1933.

Halle, Mathematisches Institut der Universität.

(Eingegangen am 26. Februar 1955.)

TAUSSKY, O.
Math. Zeitschr. Bd. 63, S. 286—289 (1955)

Unimodular integral circulants*.

In memoriam ISSAI SCHUR.

By

OLGA TAUSSKY.

Some properties of the "discriminant matrix" $(\alpha_i^{(S_k)})$ of a normal algebraic number field of degree n were investigated in two previous notes (1,2). Here the α_i form an integral basis of the field and the S_k are the elements of the GALOIS group[1]). In the special case when the α_i form a normal basis various problems concerning group matrices arise, among others, questions concerning unimodular group matrices whose elements are rational integers. If the field is cyclic circulant[2]) matrices appear, i.e. matrices $C=(c_{ik})$ with $c_{ik}=c_{k-i+1}$ where the suffixes are considered mod n. In particular the following theorem was obtained which will be studied further in the present note:

Theorem 1. *Let C be a unimodular circulant of rational integers of the form AA' where A is a unimodular matrix of rational integers. Let C be an $n\times n$ matrix. Then A is the product of a circulant, a permutation matrix and a diagonal matrix whose elements are ± 1.*

Now let $n=p$ be (in particular) a prime and $B=(b_{k-i+1})$ a unimodular circulant[3]). Since B is unimodular, we have

$$b_1+\cdots+b_p=\pm 1$$

$$b_1+b_2\zeta+\cdots+b_p\zeta^{p-1}=\varepsilon_B$$

where ε_B is a unit in the field $F(\zeta)$ generated by ζ, a p-th root of unity.

We first prove two lemmas which are concerned with these units which are "associated" with unimodular circulants.

Lemma 1. *Let B be a unimodular circulant and ε_B its associated unit. Then BB' has $\varepsilon_B\bar{\varepsilon}_B$ as its associated unit.*

Proof. Since the associated unit is also an eigenvalue and since circulants are polynomials in the permutation matrix, the statement follows at once.

It is convenient to consider the product $\varepsilon\bar{\varepsilon}$ as the norm of ε with respect to the field $F(\zeta+\zeta^{-1})$.

A converse of Lemma 1 is:

---

*) This work was supported (in part) by the Office of Naval Research.

[1]) The idea of studying discriminant matrices was suggested by I. SCHUR's treatment of GAUSS sums; see E. LANDAU. Vorlesungen über Zahlentheorie, Leipzig 1927, p. 162—166.

[2]) For a detailed study of circulants see e.g. (3).

[3]) In what follows all circulants are assumed to have rational integers as elements.

Lemma 2. *Let C be a unimodular circulant and assume that its associated unit $\varepsilon_C = \varepsilon\bar{\varepsilon}$, where ε is the unit $b_1 + b_2\zeta + \cdots + b_p\zeta^{p-1}$. Then $C = BB'$ where B is the circulant whose first row is $b_1, b_2, \ldots, b_p$.*

Proof. Consider the equation

$$c_1 + c_2\zeta + \cdots + c_p\zeta^{p-1} = (b_1 + b_2\zeta + \cdots + b_p\zeta^{p-1})(b_1 + b_2\zeta^{-1} + \cdots + b_p\zeta^{-(p-1)}).$$

Since the expansion in powers of ζ is unique, we obtain by comparing coefficients

$$\begin{aligned} c_1 &= b_1^2 + b_2^2 + \cdots + b_p^2 \\ c_2 &= b_1 b_p + b_2 b_1 + \cdots + b_p b_{p-1} \\ c_3 &= b_1 b_{p-1} + b_2 b_p + \cdots + b_p b_{p-2} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ c_p &= b_1 b_2 + b_2 b_3 + \cdots + b_p b_1. \end{aligned}$$

From this $C = BB'$ follows by inspection.

From Lemmas 1, 2 and Theorem 1 we conclude:

Theorem 2. *Let $C = AA'$ be a unimodular circulant where A is a unimodular matrix with rational integers as elements. Then the unit ε_C associated with C is the norm of a unit ε of $F(\zeta)$ with respect to $F(\zeta + \zeta^{-1})$. If $\varepsilon = b_1 + b_2\zeta + \cdots + b_p\zeta^{p-1}$ then $C = BB'$ where B is the circulant whose first row is $(b_1, \ldots, b_p)$.*

We will now prove a special instance of Theorem 2 (for $p = 5$) without explicitly using the fact that the matrix is of the form AA'.

Theorem 3. *Let C be a symmetric and positive definite unimodular 5×5 circulant. Then ε_C is the norm of a unit of $F(\zeta)$ with respect to $F(\zeta + \zeta^{-1})$ where ζ is a 5-th root of unity.* (By Lemma 2 this implies that $C = BB'$ where B is a circulant derived from the unit.)

Proof of Theorem 3[4]).

Let c_1, c_2, c_3, c_4, c_5 be the elements of the first row of C. Since C is symmetric, we have

$$c_2 = c_5, \quad c_3 = c_4.$$

We have therefore the relation

$$c_1 + 2c_2 + 2c_3 = \pm 1.$$

The value -1 is not possible since the determinant of C is $+1$ by assumption and since its value is the product of $c_1 + 2c_2 + 2c_3$ and the norm of a unit in $F(\zeta)$. This norm must have the value $+1$ since the field is complex. We further know that

$$c_1 + c_2\zeta + c_3\zeta^2 + c_3\zeta^3 + c_2\zeta^4 = \varepsilon$$

where ε is a unit in the field $F(\zeta)$. From the special form of ε it is evident that it lies in the quadratic subfield generated by $\eta = \zeta + \zeta^{-1}$ which is the

[4]) I am indebted to J. Todd for some helpful suggestions.

field generated by $\sqrt{5}$. We now show that ε is the norm of a unit of $F(\zeta)$ with respect to $F(\eta)$. For this purpose we show that ε is the square of another unit in $F(\eta)$. The field $F(\eta)$ has η as a fundamental unit. What we are going to prove is that

$$\varepsilon = c_1 - c_3 + (c_2 - c_3)\eta. \tag{1}$$

is an even power of η under the assumption that

$$c_1 + 2c_2 + 2c_3 = 1. \tag{2}$$

We start by showing some simple facts about the powers of η. The number η satisfies the equation

$$x^2 + x - 1 = 0.$$

From this it follows easily (e.g. by induction) that in

$$\eta^n = x_n + y_n\eta$$

the product

$$x_n y_n < 0, \quad n > 1, \tag{3}$$

in particular:

$$\begin{aligned} &x_n > 0,\ y_n < 0,\ n \text{ even}, > 0 \\ &x_n < 0,\ y_n > 0,\ n \text{ odd}, > 1. \end{aligned} \tag{4}$$

From the positive definiteness of C it follows that

$$c_1 > 0. \tag{5}$$

We next show that

$$c_1 - c_3 > 0. \tag{6}$$

For, if $c_1 - c_3 < 0$, then $c_3 > 0$, hence from (2) we have $c_2 < 0$, which implies $c_2 - c_3 < 0$. This would imply that the coordinates in (1) are of the same sign which contradicts (3). Hence (6) holds because $c_1 - c_3 = 0$ implies $c_2 - c_3 = \pm 1$ in view of (1). This, however, contradicts (2).

From (4) it follows that (1) is an even power of η or the negative of an odd power. We now show that the latter is not true: for, it follows from (2) that

$$c_1 + 2c_2 + 2c_3 = (c_1 - c_3) + 2(c_2 - c_3) + 5c_3 = 1,$$

hence

$$c_1 - c_3 + 2(c_2 - c_3) \equiv 1\,(5). \tag{7}$$

So we have to show that

$$x_n + 2y_n \equiv 1\,(5) \tag{8}$$

cannot happen for $-\eta^n$, n odd.

We will even show that $x_n + 2y_n \equiv \pm 1 (5)$ for n odd, $n > 1$. This can again be shown by induction: for $n = 3$ we have

$$\eta^3 = -1 + 2\eta,$$

hence the assertion is true for $n = 3$. Putting again

$$\eta^n = x_n + y_n \eta,$$

it follows that

$$x_{2n+3} + 2y_{2n+3} = -x_{2n+1} + 3y_{2n+1} \equiv -(x_{2n+1} + 2y_{2n+1}) (5),$$

hence (8) cannot happen for n odd.

We have now shown that ε is an even power of η, hence it is the norm of a unit in $F(\eta)$.

References.

[1] TAUSSKY, O.: Normal matrices in some problems in algebraic number theory. Proc. Internat. Congr. Math., Amsterdam, 1954. — [2] NEWMAN, M., and O. TAUSSKY: On a generalization of the normal basis in abelian algebraic number fields. To appear in Comm. Pure Appl. Math. — [3] HAMBURGER, H. L., and M. E. GRIMSHAW: Linear Transformations, Cambridge 1951, pp. 94—96.

National Bureau of Standards, Washington, D. C. (USA).

(Eingegangen am 7. März 1955)

HIRSCH, K. A.
Math. Zeitschr. Bd. 63, S. 290—294 (1955)

Über lokal-nilpotente Gruppen.

ISSAI SCHUR zum Gedächtnis.

Von

KURT A. HIRSCH.

An anderer Stelle ([8], [9]) habe ich den folgenden Satz bewiesen:

Satz 1. *Eine Gruppe G mit Maximalbedingung für Untergruppen, in der jede echte Untergruppe von ihrem Normalisator verschieden ist, ist nilpotent*[1]).

Es liegt nahe, zu fragen, ob die Voraussetzung der Maximalbedingung abgeschwächt oder ganz fortgelassen werden kann, wenn die Folgerung entsprechend abgeschwächt wird. In dieser Richtung hat PLOTKIN den folgenden Satz bewiesen ([11]; s. auch [10], § 62):

Satz 2. *Eine Gruppe G, in der jede echte Untergruppe von ihrem Normalisator verschieden ist, ist lokal-nilpotent*[2]).

Ich will nun hier zeigen, daß eine verschärfte Form der Schlüsse, die ich bei dem Beweis von Satz 1 benutzt habe, ausreicht, um einen einfachen Beweis von Satz 2 zu erbringen. Zu diesem Zweck brauche ich zwei Sätze, die auch an sich Interesse beanspruchen dürfen.

Satz 3. *Jede beliebige Gruppe G besitzt einen eindeutig bestimmten maximalen lokal-nilpotenten Normalteiler, der alle lokal-nilpotenten Normalteiler von G umfaßt.*

Dieser Satz ist eine unmittelbare Folge von

Satz 4. *Das Erzeugnis zweier lokal-nilpotenter Normalteiler A und B einer beliebigen Gruppe G ist selbst lokal-nilpotent.*

Satz 4 und Satz 3 entsprechen den beim Beweis von Satz 1 benutzten Tatsachen, daß das Erzeugnis zweier nilpotenter Normalteiler einer beliebigen Gruppe selbst ein nilpotenter Normalteiler ist, und daß daher jede Gruppe mit Maximalbedingung für Normalteiler einen eindeutigen maximalen nilpotenten Normalteiler besitzt, der alle nilpotenten Normalteiler umfaßt ([7], Theorem 3.1; s. auch [2], Satz 14).

Schließlich will ich noch auf ein mehrfach zu benutzendes Korollar eines Satzes von BAER hinweisen ([1], S. 363; s. auch [10], § 63):

Satz 5. *Jede endlich erzeugte nilpotente Gruppe erfüllt die Maximalbedingung für Untergruppen.*

[1]) Eine Gruppe heißt nilpotent, wenn sie eine Zentralreihe endlicher Länge besitzt.

[2]) Lokal-nilpotent heißt eine Gruppe, in der jedes endliche System von Elementen eine nilpotente Untergruppe erzeugt.

Beweis von Satz 4. 1. A und B seien zwei lokal-nilpotente Normalteiler einer Gruppe G. Wenn a und b zwei beliebige Elemente von A und B sind, so liegt der Kommutator $a' = [a, b] = b'$ sowohl in A als auch in B. Nun ist die Gruppe $\{b', b\}$ als endlich erzeugte Untergruppe von B nach Voraussetzung nilpotent, so daß wir bei genügender Wiederholung der Kommutatorbildung $[b', b, b, \ldots, b] = 1$ erhalten. Wegen $b' = a'$ besagt dies: Setzen wir $[a, b] = a'$, $[a', b] = a''$ und so weiter, so bricht die Folge $a, a', a'', \ldots$ nach endlich vielen Schritten mit dem Einheitselement ab.

2. Um zu zeigen, daß $\{A, B\}$ lokal-nilpotent ist, nehmen wir in $\{A, B\}$ ein beliebiges endliches System von Elementen, das wir in der Form $a_1 b^{(1)}$, $a_2 b^{(2)}, \ldots, a_n b^{(n)}$ schreiben wollen. Wir werden die Gruppe $\{a_1 b^{(1)}, a_2 b^{(2)}, \ldots, a_n b^{(n)}\}$ dadurch als nilpotent erweisen, daß wir sie in eine nilpotente Untergruppe von $\{A, B\}$ einbetten. Wir setzen $A_0 = \{a_1, a_2, \ldots, a_n\}$ und $B^* = \{b^{(1)}, b^{(2)}, \ldots, b^{(n)}\}$. Als endlich erzeugte Untergruppe von B ist B^* nilpotent und erfüllt nach Satz 5 die Maximalbedingung für Untergruppen. Also hat B^* (s. [6], Theorem 2.311) eine „Hauptreihe"

$$\{1\} = B_0 < B_1 < B_2 < \cdots < B_k = B^*, \tag{1}$$

in der die Gruppen B_i $(i = 1, 2, \ldots, k)$ Normalteiler von B^* sind und B_{i+1}/B_i eine zyklische Gruppe (von endlicher oder unendlicher Ordnung) ist. Es sei b_i, $i = 1, 2, \ldots, k$, ein Vertreter eines erzeugenden Elementes von B_i/B_{i-1} in B^*, so daß insbesondere $\{b_1, b_2, \ldots, b_i\} = B_i$ ist.

3. Wir konstruieren jetzt für jedes i $(i = 1, 2, \ldots, k)$ eine Gruppe A_i, die den folgenden Bedingungen genügt:

(i) A_i ist eine endlich erzeugte Untergruppe von A, die A_0 umfaßt.

(ii) In der aufsteigenden Folge

$$A_i < \{A_i, B_1\} < \{A_i, B_2\} < \cdots < \{A_i, B_i\} \tag{2}$$

ist jedes Glied Normalteiler des darauffolgenden, und alle Glieder sind nilpotent.

Wir beginnen mit $i = 1$. Wenn wir den unter 1. beschriebenen Prozeß der wiederholten Kommutatorbildung mit b_1 auf alle erzeugenden Elemente von A_0 anwenden, erhalten wir

$$a_1, a_1', a_1'', \ldots; \quad a_2, a_2', a_2'', \ldots; \quad \ldots; \quad a_n, a_n', a_n'', \ldots. \tag{3}$$

A_1 sei die von all diesen endlich vielen Elementen erzeugte Gruppe. Als endlich erzeugte Untergruppe von A ist A_1 nilpotent. Ferner ist A_1 Normalteiler von $\{A_1, B_1\}$, da ja für jedes Element a_m, $m = 1, 2, \ldots, n$

$$b_1^{-1} a_m^{(j)} b_1 = a_m^{(j)} a_m^{(j+1)} \in A_1 \tag{4}$$

ist. Wir zeigen nun, daß $\{A_1, B_1\}$ wiederum nilpotent ist. Als nilpotente Gruppe hat A_1 ein nicht-triviales Zentrum Z. Sei $z \neq 1$ ein beliebiges Element von Z. Wenn nicht schon $[z, b_1] = 1$ ist, so formen wir, wie oben, durch wiederholte Kommutatorbildung $z, z', z'', \ldots$ und kommen nach endlich vielen

Schritten zu einem Zentrumselement von A_1, das auch mit b_1 vertauschbar ist, also zum Zentrum von $\{A_1, B_1\}$ gehört. Wir bilden nun die Faktorgruppe nach diesem Zentrum und sehen durch vollständige Induktion, daß $\{A_1, B_1\}$ nilpotent ist.

4. Im allgemeinen Fall sei die Gruppe A_i folgendermaßen definiert. Wir nehmen die erzeugenden Elemente a_m $(m = 1, 2, \ldots, n)$ von A_0 und fügen alle diejenigen Elemente von A hinzu, die durch wiederholte Kommutatorbildung mit b_i entstehen; dann weiter die wiederholten Kommutatoren all dieser Elemente mit b_{i-1}; in dieser Weise fahren wir mit $b_{i-2}, b_{i-3}, \ldots, b_1$ fort. Die erzeugenden Elemente von A_i sind also die a_m $(m = 1, \ldots, n)$ und alle Kommutatoren der Form

$$a = [a_m, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_s}] \tag{5}$$

wobei $i \geq \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_s \geq 1$ ist.

Auf Grund der Bemerkung unter 1. gibt es nur endlich viele verschiedene Kommutatoren dieser Form, so daß die Forderung (i) für A_i erfüllt ist.

Daß nun A_i Normalteiler von $\{A_i, B_1\}$ und die letztere Gruppe nilpotent ist, sehen wir genau wie unter 3. Nehmen wir schon als bewiesen an, daß die Forderung (ii) in der Folge (2) bis zu $\{A_i, B_{j-1}\}$ erfüllt ist. Dann haben wir zu zeigen, daß $\{A_i, B_{j-1}\}$ Normalteiler von $\{A_i, B_j\} = \{A_i, B_{j-1}, b_j\}$ ist. Nun ist B_{j-1} Normalteiler von B_j. Es genügt also zu zeigen, daß für jeden Kommutator (5)

$$b_j^{-1} a\, b_j \in \{A_i, B_{j-1}\} \tag{6}$$

gilt.

Es sei r so bestimmt, daß $\alpha_r \geq j > \alpha_{r+1}$ ist. Dann setzen wir

$$[a_m, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_r}] = \bar{a},$$

wobei $\bar{a}$ wieder ein erzeugendes Element von A_i ist. Wir haben also

$$b_j^{-1} a\, b_j = b_j^{-1} [\bar{a}, b_{\alpha_{r+1}}, \ldots, b_{\alpha_s}]\, b_j = [b_j^{-1} \bar{a}\, b_j, b_j^{-1} b_{\alpha_{r+1}} b_j, \ldots, b_j^{-1} b_{\alpha_s} b_j]. \tag{7}$$

Hierin ist $b_j^{-1} \bar{a}\, b_j = \bar{a}\, [\bar{a}, b_j]$ ein Produkt von zwei Erzeugenden von A_i, und alle übrigen Elemente der Klammer liegen in B_{j-1}, da $\alpha_{r+1} \leq j - 1$ und B_{j-1} Normalteiler von B_j ist. Damit ist (6) bewiesen. Tatsächlich läßt sich aber viel mehr zeigen, nämlich daß A_i Normalteiler in $\{A_i, B_j\}$ ist. Denn wir können mit Hilfe der fundamentalen Kommutator-Identitäten

$$\begin{cases} [a\,b, c] = b^{-1} [a, c]\, b\, [b, c] = [a, c]\, [a, c]\, [a, c, b]\, [b, c] \\ [a, b\,c] = [a, c]\, c^{-1} [a, b]\, c = [a, c]\, [a, b]\, [a, b, c] \end{cases} \tag{8}$$

die rechte Seite von (7) sukzessiv so umformen, daß ein Produkt von Elementen der Form (5) entsteht. Hierzu stellen wir für jedes $\alpha \leq \alpha_{r-1}$ das Element $b_j^{-1} b_\alpha b_j$ in der Form

$$b_j^{-1} b_\alpha b_j = b_\alpha b_\beta b_\gamma \ldots b_\omega \tag{9}$$

dar, mit $\alpha \geq \beta \geq \gamma \geq \cdots \geq \omega \geq 1$, was in einer nilpotenten Gruppe stets möglich ist. Die etwas langwierigen Einzelheiten des Beweises brauchen hier nicht durchgeführt zu werden.

Die Nilpotenz von $\{A_i, B_j\}$ ergibt sich aus der von $\{A_j, B_{j-1}\}$ genau wie oben unter 3. Wir sehen sogar, daß das Zentrum von $\{A_i, B_j\}$ Elemente des Zentrums von A_i enthält. Schließlich haben wir in der Gruppe $\{A_k, B_k\} = \{A_k, B^*\}$ die gesuchte nilpotente Gruppe, die $\{A_0, B^*\}$ und daher

$$\{a_1 b^{(1)}, a_2 b^{(2)}, \ldots, a_n b^{(n)}\}$$

als Untergruppe enthält. Damit ist Satz 4 bewiesen.

Bemerkung. Wenn die Vermutung zutrifft, daß das Produkt zweier nilpotenter Gruppen auflösbar ist, dann läßt sich Satz 4 in wenigen Zeilen beweisen. Es gilt nämlich nach GRUENBERG [3], daß eine auflösbare, endlich erzeugte Gruppe G nilpotent ist, wenn für je zwei beliebige Elemente x, y von G eine Kommutator-Identität $[x, y, y, \ldots, y] = 1$ besteht.

Beweis von Satz 3. Das Erzeugnis einer beliebigen Menge A_α von lokal-nilpotenten Normalteilern einer Gruppe ist selbst ein lokal-nilpotenter Normalteiler. Denn jedes Element $a \in A$ ist ein endliches Produkt $a_{\alpha_1} a_{\alpha_2} \ldots a_{\alpha_k}$, $a_{\alpha_i} \in A_{\alpha_i}$, also ein Element der Gruppe $\{A_{\alpha_1}, A_{\alpha_2}, \ldots, A_{\alpha_k}\}$, die nach Satz 4 lokal-nilpotent ist. Jedes endliche System $a_1, a_2, \ldots, a_n$ von Elementen von A liegt daher in dem von einer endlichen Menge von A_α's erzeugten Normalteiler; daher ist $\{a_1, \ldots, a_n\}$ nilpotent, was zu beweisen war.

Satz 2 folgt nun in wenigen Schritten, wie beim Beweis von Satz 1.

Beweis von Satz 2. Sei G eine beliebige Gruppe. Jedes Gruppenelement $g \in G$ liegt in einer lokal-nilpotenten, sogar nilpotenten Untergruppe, z.B. der von g erzeugten zyklischen Gruppe. Die Vereinigungsmenge einer beliebigen aufsteigenden Folge von lokal-nilpotenten Untergruppen ist wieder lokal-nilpotent. Daher gibt es nach dem ZORNschen Lemma in jeder beliebigen Gruppe maximale lokal-nilpotente Untergruppen. Sei M eine solche und $N(M)$ ihr Normalisator. Dann ist $N(M)$ sein eigner Normalisator. Denn gäbe es ein Gruppenelement g, das mit $N(M)$ vertauschbar wäre, ohne in $N(M)$ zu liegen, so würde M durch g in eine Gruppe $M' = g^{-1} M g \neq M$ transformiert, die ebenfalls Normalteiler von $N(M)$ und lokal-nilpotent wäre. Dann wäre aber $\{M, M'\}$ nach Satz 4 lokal-nilpotent, im Widerspruch zu der Maximaleigenschaft von M.

Wenn nun G eine Gruppe ist, in der jede echte Untergruppe von ihrem Normalisator verschieden ist, so sehen wir, daß $N(M) = G$, also daß M ein Normalteiler von G ist. Dann umfaßt aber der maximale lokal-nilpotente Normalteiler (Satz 3) alle Elemente von G, d.h., G ist lokal-nilpotent.

Beweis von Satz 5. Sei G eine endlich erzeugte nilpotente Gruppe. Die Kommutator-Faktorgruppe G/G' ist als homomorphes Bild von G ebenfalls endlich erzeugbar. Nach einem Satz von BAER ([1], S. 363), den man auch durch Induktion nach der Klasse von G leicht direkt beweisen kann,

sind dann alle Faktorgruppen G^i/G^{i+1} der unteren Zentralreihe von G endlich erzeugbar, und daher erfüllt G die Maximalbedingung für Untergruppen ([5], Theorem 1.3)[3]).

Ich möchte zum Schluß die Gelegenheit benutzen, einen Irrtum in [9] richtigzustellen, auf den mich G. HIGMAN aufmerksam gemacht hat. Es ist nicht richtig, daß in der von mir dort angegebenen nicht-nilpotenten Gruppe alle maximalen Untergruppen Normalteiler sind. Man kann aber ein Beispiel einer solchen Gruppe herstellen, indem man eine der von HIGMAN und NEUMANN [4] konstruierten nicht-nilpotenten Gruppen ohne maximale Untergruppen mit einer zyklischen Gruppe von Primzahlordnung direkt multipliziert.

Literatur.

[1] BAER, R.: Representations of Groups as Quotient Groups. II. Trans. Amer. Math. Soc. **58**, 348—389 (1945). — [2] FITTING, H.: Beiträge zur Theorie der Gruppen endlicher Ordnung. Jber. dtsch. Math.-Ver. **48**, 77—141 (1938). — [3] GRUENBERG, K.W.: Two Theorems on Engel Groups. Proc. Cambr. Phil. Soc. **49**, 377—380 (1953). — [4] HIGMAN, G., and B. H. NEUMANN: On two Questions of N. Itô. J. London Math. Soc. **29**, 84—88 (1954). — [5] HIRSCH, K. A.: On Infinite Soluble Groups. I. Proc. London Math. Soc. **44**, 53—60 (1938). — [6] HIRSCH, K. A.: On Infinite Soluble Groups. II. Proc. London Math. Soc. **44**, 336—344 (1938). — [7] HIRSCH, K. A.: On Infinite Soluble Groups. III. Proc. London Math. Soc. **49**, 184—194 (1947). — [8] HIRSCH, K. A.: A Characteristic Property of Nilpotent Groups. Proc. Int. Congr. of Math. Cambridge (Mass.) 1950. — [9] HIRSCH, K. A.: Eine kennzeichnende Eigenschaft nilpotenter Gruppen. Math. Nachr. **4**, 47—49 (1951). — [10] KUROŠ, A. G.: Teoriya Grupp, 2. (russische) Aufl. Moskau 1953. — [11] PLOTKIN, B. I.: Zur Theorie der lokal-nilpotenten Gruppen (russisch). Dokl. Akad. Nauk USSR. **76**, 639—641 (1951).

[3]) Andere Beweise von Satz 5 sind kürzlich von P. HALL [Proc. London Math. Soc. (3) **4**, 419—436 (1954)] und von S. A. JENNINGS [Canadian J. of Math. **7**, 169—187 (1955)] veröffentlicht worden [bei der Korrektur hinzugefügt].

Queen Mary College, University of London, London E 1 (England).

(Eingegangen am 4. März 1955.)

SKOLEM, TH.
Math. Zeitschr. Bd. 63, S. 295—312 (1955)

Einige Bemerkungen über die Auffindung der rationalen Punkte auf gewissen algebraischen Gebilden.

ISSAI SCHUR zum Gedächtnis.

Von

TH. SKOLEM.

Einleitung.

Bekanntlich hat die Frage nach den rationalen Punkten auf einer ebenen algebraischen Kurve eine ziemlich erschöpfende Antwort gefunden in dem Falle einer Kurve vom Geschlecht Null, zum Teil auch für Kurven vom Geschlecht Eins. Meine erste Bemerkung in diesem Aufsatz betrifft die Ausdehnung eines bekannten Satzes über unicursale ebene Kurven auf solche in Räumen beliebiger Dimensionszahl. Ich mache auch eine Bemerkung über Parameterdarstellungen durch gewisse transzendente Funktionen derart, daß die Argumente der rationalen Punkte einen Modul bilden.

Was algebraische Flächen betrifft, so ist es schon schwierig genug, die rationalen zu studieren, d. h. diejenigen, für welche eine Parameterdarstellung mittels rationaler Funktionen möglich ist. Sieht man von den ziemlich trivialen Flächen zweiten Grades ab, so erweist sich bereits das Studium der Flächen dritten Grades im dreidimensionalen Raume als sehr weitläufig. Ihre Behandlung ist oft ziemlich erkünstelt und unsystematisch ausgeführt worden, und meine Bemerkungen im folgenden haben den Zweck, natürlichere Verfahren zu entwickeln.

§ 1.

Ich betrachte Kurven, deren Gleichungen rationale Koeffizienten haben. Ich benutze als eine bekannte geometrische Tatsache, auf deren Beweis ich nicht näher eingehe, daß man das Koordinatensystem so wählen kann, daß die Projektion der betrachteten Kurve in eine der Koordinatenebenen nur endlich viele Punkte besitzt, die gleichzeitig Projektion von mehreren Punkten der Raumkurve sind. Dies bedeutet, daß die Gleichungen der betrachteten Kurve C die Form

$$f(x_1, x_2) = 0, \qquad x_i = g_i(x_1, x_2)/h_i(x_1, x_2), \quad i = 3, \ldots, n$$

haben, wo f und die g_i sowie h_i Polynome mit rationalen Koeffizienten sind. Jeder rationale Punkt von C projiziert sich dann in einen rationalen Punkt der Kurve $f(x_1, x_2) = 0$, welche ich $C_{1,2}$ nenne. Umgekehrt ist jeder rationale Punkt auf $C_{1,2}$ die Projektion eines solchen auf C, wenn eventuell von den

endlich vielen Punkten abgesehen wird, wo für ein i die beiden Polynome g_i und h_i gleichzeitig verschwinden. Durch diese Bemerkung wird klar, daß die Aufgabe, die rationalen Punkte auf einer Raumkurve C zu finden, im wesentlichen auf dieselbe Aufgabe für die Projektionskurve $C_{1,2}$ zurückführbar ist.

Falls nun C unicursal ist, so ist $C_{1,2}$ unicursal, und man kann entscheiden, ob rationale Punkte auf $C_{1,2}$ existieren. Wie üblich nenne ich einen Punkt P gewöhnlich, wenn es darin eine eindeutig bestimmte Tangente gibt, d.h. wenn die Sekante durch zwei benachbarte Punkte P_1 und P_2 eine einzige Grenzstellung besitzt, wenn P_1 und P_2 unabhängig voneinander gegen P streben. Die sonstigen Punkte heißen singulär. Dann hat $C_{1,2}$ entweder keinen gewöhnlichen rationalen Punkt oder unendlich viele, und man kann entscheiden, ob das eine oder das andere eintritt.

Ist nun P ein gewöhnlicher rationaler Punkt auf C, so können wir durch geeignete Wahl des Koordinatensystems erreichen, daß P sich auf $C_{1,2}$ in einen gewöhnlichen rationalen Punkt projiziert. Die unendlich vielen sonst vorhandenen rationalen gewöhnlichen Punkte auf $C_{1,2}$ sind dann die Projektionen von unendlich vielen gewöhnlichen rationalen Punkten auf C. Hierdurch sieht man ein, daß jede unicursale Raumkurve entweder keinen gewöhnlichen rationalen Punkt hat oder deren unendlich viele; außerdem ist entscheidbar, welche Alternative realisiert ist.

Man hat aber bekanntlich auch den Satz (vgl. SKOLEM [7], S. 73):

Eine ebene Kurve vom Geschlecht Null besitzt stets und nur dann einen rationalen gewöhnlichen Punkt, wenn sie mindestens einen gewöhnlichen reellen Punkt enthält und ihre Gleichung in jedem perfekten p-adischen Körper eine Lösung hat, die einen gewöhnlichen Punkt darstellt.

Ich werde eine ganz kurze Überlegung entwickeln, durch die man einsehen kann, *daß dieser Satz auch für jede unicursale Kurve in einem Raume beliebiger Dimensionszahl gültig sein muß.*

Nehmen wir an, daß C einen gewöhnlichen Punkt besitzt in jedem der perfekten p-adischen Körper und im reellen Körper. Es ist bekannt, daß ein gewöhnlicher Punkt in einem solchen Körper nicht isoliert sein kann, so daß in der Tat innerhalb einer gewissen Umgebung von P sogar ein Kontinuum von gewöhnlichen Punkten vorhanden ist. Nennen wir jeden Punkt auf $C_{1,2}$, der entweder singulär ist oder so beschaffen, daß ein g_i gleichzeitig mit h_i verschwindet, einen Ausnahmepunkt, so ist klar, daß entweder P selbst oder ein Nachbarpunkt Q einen Projektionspunkt P' bzw. Q' auf $C_{1,2}$ hat, der kein Ausnahmepunkt ist. Da also $C_{1,2}$ in jedem der perfekten Körper einen gewöhnlichen Punkt hat, so besitzt sie nach dem erwähnten Satz einen rationalen gewöhnlichen Punkt. Wie oben bemerkt, gibt es dann sogar unendlich viele rationale gewöhnliche Punkte auf $C_{1,2}$. Jeder derselben, der kein Ausnahmepunkt ist, gibt zufolge der Gleichungen der Kurve C zu einem rationalen Punkt auf C Anlaß. Hierdurch ist bewiesen, daß der Satz auch für unicursale Raumkurven wahr ist.

§ 2.

Man kann allgemeiner die Lösbarkeit in algebraischen Zahlkörpern untersuchen. Wegen der Zurückführbarkeit auf ebene Kurven, und weil jede ebene Kurve vom Geschlecht Null entweder mit einer Geraden oder einem Kegelschnitt birational mit rationalen Koeffizienten äquivalent ist (nämlich je nachdem der Grad der Kurve ungerade oder gerade ist), so ist diese weitere Lösbarkeitsfrage für unicursale Kurven nur für einen Kegelschnitt aktuell. Man kann wohl kaum sagen, daß man eine wirkliche Übersicht über die kleinsten Körper hat, in denen Lösbarkeit besteht. Es ist aber jedenfalls klar: Wenn Lösbarkeit in einem Körper K stattfindet, der von ungeradem Relativgrad in bezug auf einen Unterkörper k ist, so besteht bereits Lösbarkeit in k. Sei nämlich $2m-1$ der Relativgrad und sei $P, P', \ldots, P^{(2m-2)}$ ein System konjugierter Punkte in bezug auf k. Dann kann man durch diese Punkte eine Kurve m-ten Grades (in der Tat unendlich viele) mit Koeffizienten in k legen; diese schneidet dann den Kegelschnitt in einem einzigen $2m$-ten Punkt Q, dessen Koordinaten dann zu k gehören müssen.

Ein gewisses Gegenstück hierzu ist, daß Körper von beliebig hohem geraden Grade existieren, worin Lösbarkeit nicht stattfindet, sogar Körper, die durch wiederholtes Ausziehen von Quadratwurzeln entstehen. Dies ist sofort klar in dem Falle, wo der Kegelschnitt keinen reellen Punkt besitzt. Sonst ist die Lösbarkeit von

$$a x^2 + b y^2 + c z^2 = 0, \tag{1}$$

wo $(a, b) = (a, c) = (b, c) = 1$ damit gleichbedeutend, daß

$$d = -ab - ac - bc$$

quadratischer Rest modulo abc ist. Nennt man jeden natürlichen Primfaktor p von abc derart, daß $\left(\frac{d}{p}\right) = -1$ ist, einen Nichtteiler, so bleibt Unlösbarkeit in einem Körper K bestehen dann und nur dann, wenn mindestens einer dieser Nichtteiler in K mindestens einen Primidealteiler ungeraden Grades hat. Ist nämlich $\mathfrak{p}$ ein solches Primideal, g sein Grad, so ist bekanntlich $\left(\frac{d, K}{\mathfrak{p}}\right)$ gleich $\left(\frac{d}{p}\right)^g$. Betrachtet man z.B. die in rationalen Zahlen $\neq 0$ unlösbare Gleichung

$$x^2 + y^2 - 3 z^2 = 0,$$

so bleibt diese unlösbar in jedem Körper, der durch Adjunktion einer beliebigen Anzahl von Quadratwurzeln $\sqrt{p}$ entsteht, wo $p \equiv 1$ oder $-1 \pmod{12}$. Denn da $\left(\frac{3}{p}\right) = +1$ ist für solche p, so zerlegt sich bekanntlich 3 in jedem dieser Körper in Primideale ersten Grades, so daß -1 quadratischer Nichtrest bleiben muß für diese Primideale. Infolgedessen bleibt die Gleichung auch unlösbar in dem Körper unendlich hohen Grades, der von sämtlichen Quadratwurzeln $\sqrt{p}$ mit $p \equiv \pm 1 \pmod{12}$ herrührt.

Was insbesondere die Lösbarkeit von (1) im quadratischen Körper $k(\sqrt{m})$ betrifft, so findet man leicht, daß sie mit der Lösbarkeit der quaternären Gleichung

$$a x^2 + b y^2 + c z^2 + m a b c u^2 = 0$$

in rationalen Zahlen gleichwertig ist. Hieraus folgt wieder unter anderem, daß $a x^2 + b y^2 + c z^2$ eine universale Form ist, d.h. jede rationale Zahl darstellt für gewisse rationale Werte von x, y, z, wenn sie eine Nullform ist, d.h. (1) rational lösbar ist. Wegen einer ähnlichen Betrachtung im folgenden ist es zweckmäßig, hier zu bemerken: *Wenn* (1) *lösbar ist, so ist die Gleichung*

$$a x^2 + b y^2 + c z^2 = r \tag{2}$$

durch lineare Ausdrücke

$$x = a_1 r + a_2, \quad y = b_1 r + b_2, \quad z = c_1 r + c_2 \tag{3}$$

identisch in r lösbar. Sind in der Tat (a_1, b_1, c_1) und (a_2', b_2', c_2') zwei verschiedene rationale Punkte auf dem Kegelschnitt (1), so ist

$$a a_1 a_2' + b b_1 b_2' + c c_1 c_2' \neq 0,$$

und folglich kann man a_2, b_2, c_2 proportional a_2', b_2', c_2' so wählen, daß (2) mittels (3) erfüllt wird. Man kann aber auch weitere Parameterlösungen in unübersehbarer Menge herleiten. Immer, wenn (a_3, b_3, c_3) ein rationaler Punkt P auf (1) ist, dagegen (a_4, b_4, c_4) ein Punkt auf der Tangente in P, dann ergibt

$$x = a_3 s + a_4, \quad y = b_3 s + b_4, \quad z = c_3 s + c_4$$

eine Lösung von

$$a x^2 + b y^2 + c z^2 = h,$$

wo h eine gewisse rationale Konstante ist. Bedeuten i, j Quaternionen derart, daß

$$i^2 = -ab, \quad j^2 = -ac, \quad ij + ji = 0,$$

und bildet man das Produkt

$$\begin{aligned}(a(a_3 s + a_4) + (b_3 s + b_4) i + (c_3 s + c_4) j)\,(a(a_1 r + a_2) + (b_1 r + b_2) i + \\ + (c_1 r + c_2) j)\,(a(a_3 s + a_4) + (b_3 s + b_4) i + (c_3 s + c_4) j) \\ = aA + Bi + Cj,\end{aligned}$$

wo A, B, C Polynome sind, die quadratisch in bezug auf s und linear in r sind, so erkennt man, daß

$$a A^2 + b B^2 + c C^2 = a^2 h^2 r$$

identisch in s und r erfüllt ist. — Mittels derselben Methode kann man weitere Lösungen der Gl. (2) erhalten, wo x, y, z als Polynome in $r, s_1, s_2, \ldots$ ausgedrückt sind, so daß (2) identisch in allen Variablen erfüllt ist.

Bekanntlich hat man den MORDELL-WEILschen Satz, daß die rationalen Punkte auf Kurven vom Geschlecht 1 mittels elliptischer Funktionen eines Parameters gegeben sind, wobei die Argumente einen Modul bilden. Man

muß erwarten, daß etwas Analoges für die einfacheren Kurven vom Geschlecht Null stattfindet. Dies ist nun insofern in trivialer Weise der Fall, als man ja die rationalen Werte eines Parameters x durch die Gleichung $x = e^u$ ausdrücken kann, und dann bilden die zugehörigen Argumentwerte von u einen Modul. Der Unterschied ist offenbar nur, daß man für die unicursalen Kurven einen Modul mit unendlicher Basis hat; man kann also, wenn man den Begriff des Ranges für diese Kurven benutzen will, sagen, daß sie einen unendlichen Rang haben. Von etwas größerem Interesse ist aber die folgende Parameterdarstellung der Lösungen der Gleichung

$$x^2 - m y^2 = 1. \tag{4}$$

Die allgemeine Lösung ist hier durch den Ansatz

$$x + y\sqrt{m} = \frac{u + v\sqrt{m}}{u - v\sqrt{m}}$$

gegeben, wo u und v ganze Zahlen sind, nicht beide Null. Nach einem Satze, den ich vor einigen Jahren bewiesen habe [8] besitzt die multiplikative Gruppe aller Zahlen $\neq 0$ eines algebraischen Körpers stets eine, allerdings unendliche, Basis. Dies ist trivial, wenn die Anzahl der Idealklassen 1 ist, ist aber auch sonst wahr. Man hat deshalb

$$u + v\sqrt{m} = \pi_1^{x_1} \pi_2^{x_2} \dots,$$

wo die π_ν Basiselemente sind. Hieraus folgt

$$x + y\sqrt{m} = \alpha_1^{x_1} \alpha_2^{x_2} \dots. \tag{5}$$

wobei $\alpha_\nu = \frac{\pi_\nu}{\pi_\nu'}$. Ich illustriere die Anwendung dieser Formel (5) an einigen typischen Beispielen.

Beispiel 1. Sämtliche rationalen Punkte auf dem Kreise

$$x^2 + y^2 = 1$$

sind gegeben durch die Formeln

$$x = \cos(\alpha_1 x_1 + \alpha_2 x_2 + \cdots), \quad y = \sin(\alpha_1 x_1 + \alpha_2 x_2 + \cdots).$$

Dabei sind die x_ν beliebige ganze rationale Zahlen; die Zahlen α_ν sind dadurch bestimmt, daß

$$\cos\alpha_\nu + i \sin\alpha_\nu = \frac{\pi_\nu^2}{p_\nu}$$

ist, wo π_ν die Faktoren $a + bi$ mit $a > 0$, $b > 0$ aller natürlicher Primzahlen p_ν der Form $4h + 1$ und der Zahl 2 durchläuft.

Beispiel 2. Sämtliche rationalen Punkte auf der Kurve

$$x^2 + 5y^2 = 1$$

sind gegeben durch

$$x = \cos(\alpha_1 x_1 + \alpha_2 x_2 + \cdots), \quad y = \frac{1}{\sqrt{5}} \sin(\alpha_1 x_1 + \alpha_2 x_2 + \cdots),$$

wenn jetzt die α_r erstens die Zahlen $\frac{\pi_r^2}{p_r}$ durchlaufen, wo π_r den Faktor $a + b\sqrt{-5}$ mit $a > 0$, $b > 0$ der natürlichen Primzahlen $\equiv 1,9 \pmod{20}$ bedeutet, und wenn die α_r zweitens alle Zahlen durchlaufen, wo π_r den Faktor derselben Art der Zahlen $2p$ bedeutet, p eine natürliche Primzahl $\equiv 3,7 \pmod{20}$. Denn -1, 2, die natürlichen Primzahlen $\equiv 11, 13, 17, 19 \pmod{20}$, die Zahlen $a + b\sqrt{-5}$, $a > 0$, $b \neq 0$, welche Faktoren der Primzahlen $\equiv 1,9 \pmod{20}$ sind, bilden zusammen mit den Zahlen $a + b\sqrt{-5}$, $a > 0$, $b \neq 0$, welche Faktoren von $2p$ mit $p \equiv 3,7 \pmod{20}$ sind, eine Basis der multiplikativen Gruppe aller Zahlen $\neq 0$ des Körpers $k(\sqrt{-5})$.

Beispiel 3. Für $m = 2$ sind die rationalen Lösungen von (4) gegeben durch

$$x = \cosh(\alpha_1 x_1 + \alpha_2 x_2 + \cdots), \quad y = \frac{1}{\sqrt{2}} \sinh(\alpha_1 x_1 + \alpha_2 x_2 + \cdots),$$

wobei etwa

$$\cosh \alpha_1 = -1, \quad \sinh \alpha_1 = 0; \quad \cosh \alpha_2 = 3, \quad \sinh \alpha_2 = 2\sqrt{2},$$

während die folgenden α_r der Reihe nach so bestimmt sind, daß

$$\cosh \alpha_r + \sinh \alpha_r = \frac{\pi_r^2}{p_r}, \quad p_r = \text{Norm}\, \pi_r = \pi_r \pi_r', \quad p_r \equiv \pm 1 \pmod{8}.$$

Für $m > 0$ ist auch die Gleichung

$$x^2 - m y^2 = -1$$

rational lösbar, wenn -1 quadratischer Rest mod m ist. Diese Lösungen können in derselben Form

$$x = \cosh A, \quad y = \frac{1}{\sqrt{m}} \sinh A$$

dargestellt werden; hierin durchläuft aber jetzt A die Elemente eines Nebenmoduls, nämlich $A = \frac{\pi}{2} i + B$, wo B die Elemente des Moduls durchläuft derart, daß

$$x = \cosh B, \quad y = \frac{1}{\sqrt{m}} \sinh B$$

die rationalen Lösungen von (4) liefern. Übrigens kann man auch sagen, daß

$$x = \sinh B, \quad y = \frac{1}{\sqrt{m}} \cosh B,$$

die Lösungen von $x^2 - m y^2 = -1$ darstellen.

Für die allgemeinere Gleichung

$$a x^2 + b y^2 = c \tag{6}$$

können die rationalen Lösungen dargestellt werden mittels Argumentwerten aus einem Nebenmodul. Ist nämlich x_0, y_0 eine spezielle Lösung, so sind die anderen gegeben mittels der Gleichung

$$x\sqrt{a} + y\sqrt{-b} = (x_0\sqrt{a} + y_0\sqrt{-b})(u + \sqrt{-ab}\,v),$$

wo

$$u^2 + abv^2 = 1. \tag{7}$$

Man findet hieraus als allgemeine Lösung von (6) im Falle $ab > 0$

$$x = \sqrt{\frac{c}{a}}\cos(\alpha + A), \quad y = \sqrt{\frac{c}{b}}\sin(\alpha + A);$$

dabei ist α so zu bestimmen, daß

$$x_0 = \sqrt{\frac{c}{a}}\cos\alpha, \quad y_0 = \sqrt{\frac{c}{b}}\sin\alpha$$

ist, während A alle Zahlen eines Moduls durchläuft, so daß

$$u = \cos A, \quad v = \frac{1}{\sqrt{ab}}\sin A$$

alle Lösungen von (7) liefern. Für $ab < 0$ hat man ähnliche Gleichungen mit cosh und sinh anstatt cos und sin.

Ich möchte noch ein Beispiel erwähnen, nämlich eine Parameterdarstellung der rationalen Punkte auf der unicursalen kubischen Kurve, deren Gleichung in der Form

$$y^2 = x^2(x + a)$$

angenommen wird. Ist $a < 0$, so hat man mit α als Parameter

$$x = -a/\sin^2\alpha, \quad y = -a\sqrt{-a}\cos\alpha/\sin^3\alpha.$$

Man erkennt, daß $\cos 2\alpha$ und $\sqrt{-a}\sin 2\alpha$ rational werden, und daß also für die rationalen Punkte α einen Modul durchläuft, der aus den Hälften der Zahlen besteht, welche die Werte des Parameters ausmachen, die nach der obigen Ausführung die Lösungen der Gleichung

$$x^2 - 1/a\,y^2 = 1$$

liefern. Ist $a > 0$, bekommt man in analoger Weise

$$x = a/\sinh^2\alpha, \quad y = (a\sqrt{a}/\sinh^3\alpha)\cdot\cosh\alpha$$

für die Zweige der Kurve zwischen dem Doppelpunkt und ∞ und für das Oval

$$x = a/\cosh^2\alpha, \quad y = a\sqrt{a}\sinh\alpha/\cos h^3\alpha,$$

und man findet wieder, daß die rationalen Punkte durch die Werte von α in einem Modul gegeben sind.

§ 3.

In diesem Paragraphen werde ich die rationalen Punkte auf kubischen Flächen betrachten, die singuläre Punkte enthalten, aber keine Regelflächen sind (vgl. SEGRE [4], [5]). Bekanntlich können höchstens vier Doppelpunkte vorkommen. Da ich natürlich voraussetze, daß die Gleichung der Fläche rationale Koeffizienten hat, so muß, falls nur ein Doppelpunkt vorkommt, dieser rational sein, falls zwei vorkommen, müssen sie entweder beide rational sein, oder sie sind konjugiert quadratisch irrational, usw. Infolgedessen ist es klar, daß es genügt, die folgenden vier Fälle zu behandeln:

1. Ein rationaler Doppelpunkt.
2. Zwei konjugiert quadratisch irrationale Doppelpunkte.
3. Drei konjugiert kubisch irrationale Doppelpunkte.
4. Vier konjugiert biquadratisch irrationale Doppelpunkte.

Nach dem Grade der Schwierigkeit werde ich diese Fälle in der Reihenfolge 1., 4., 3., 2. behandeln.

1. Ist P der rationale Doppelpunkt, so schneidet jede rationale Gerade durch P die Fläche in einen rationalen Punkt Q; und ist umgekehrt Q ein beliebiger rationaler Punkt der Fläche, so ist die Verbindungslinie zwischen P und Q rational. Man bekommt hierdurch in trivialer Weise die Koordinaten der rationalen Punkte rational ausgedrückt durch die Winkelkoeffizienten k und l der Geraden durch P.

Ein Sonderfall hiervon ist, daß eine der Unbekannten der homogen angenommenen Gleichung, etwa u, nur linear auftritt; dann kann man u als (gebrochene) rationale Funktion von x, y, z ausdrücken.

4. Die vier konjugierten Doppelpunkte seien P_i mit Koordinaten $x^{(i)}$, $y^{(i)}$, $z^{(i)}$, $u^{(i)}$, und es sei $\xi_j = 0$ die Ebene, welche alle P_i außer P_j enthält. Dann ist z.B. ξ_1 die Determinante

$$\begin{vmatrix} x^{(2)} & y^{(2)} & z^{(2)} & u^{(2)} \\ x^{(3)} & y^{(3)} & z^{(3)} & u^{(3)} \\ x^{(4)} & y^{(4)} & z^{(4)} & u^{(4)} \\ x & y & z & u \end{vmatrix},$$

also mit konjugierten $\alpha^{(j)}(\beta^{(j)}, \gamma^{(j)}, \delta^{(j)})$

$$\xi_j = \alpha^{(j)} x + \beta^{(j)} y + \gamma^{(j)} z + \delta^{(j)} u.$$

Man sieht, daß dann umgekehrt x, y, z, u durch $\xi_1, \xi_2, \xi_3, \xi_4$ so ausgedrückt sind:

$$(8) \quad \begin{cases} x = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \alpha_3 \xi_3 + \alpha_4 \xi_4, & y = \beta_1 \xi_1 + \beta_2 \xi_2 + \beta_3 \xi_3 + \beta_4 \xi_4, \\ z = \gamma_1 \xi_1 + \gamma_2 \xi_2 + \gamma_3 \xi_3 + \gamma_4 \xi_4, & u = \delta_1 \xi_1 + \delta_2 \xi_2 + \delta_3 \xi_3 + \delta_4 \xi_4, \end{cases}$$

wo die α (und ebenso β, γ, δ) für sich konjugiert in den betreffenden biquadratischen Körpern sind. Durch Einsetzen von $\xi_1, \xi_2, \xi_3, \xi_4$ statt x, y, z, u muß die Gleichung der Fläche bekanntlich die Form annehmen

$$\varepsilon_1 \xi_2 \xi_3 \xi_4 + \varepsilon_2 \xi_1 \xi_3 \xi_4 + \varepsilon_3 \xi_1 \xi_2 \xi_4 + \varepsilon_4 \xi_1 \xi_2 \xi_3 = 0,$$

und dabei müssen augenscheinlich hier $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ konjugiert werden. Diese Gleichung kann einfacher

$$S\frac{\varepsilon_1}{\xi_1} = 0$$

geschrieben werden, wo S die Spur bedeutet. Mittels dreier rationaler Parameter kann man nun die allgemeine Lösung der Gleichung

$$S\tau_1 = 0$$

angeben, und dann liefert

$$\xi_1 = \frac{\varepsilon_1}{\tau_1}$$

den allgemeinen Ausdruck für ξ_1. Mittels (8) findet man hieraus

$$x = S\frac{\alpha_1\varepsilon_1}{\tau_1}, \quad y = S\frac{\beta_1\varepsilon_1}{\tau_1}, \quad z = S\frac{\gamma_1\varepsilon_1}{\tau_1}, \quad u = S\frac{\delta_1\varepsilon_1}{\tau_1}$$

als eine vollständige Parameterdarstellung der rationalen Lösungen.

3. Mittels linearer Transformation mit rationalen Koeffizienten kann die Gleichung der Fläche auf eine solche Form gebracht werden, daß die drei Doppelpunkte P_1, P_2, P_3 in der Ebene $x=0$ liegen. Es seien K_1, K_2, K_3 die konjugierten kubischen Körper, wozu die Koordinaten von P_1, P_2, P_3 gehören. Weiter sei $\xi_1=0$ eine Ebene, welche P_2, P_3 enthält; es ist klar, daß ξ_1 so wählbar ist, daß ihre Koeffizienten zu K_1 gehören. Unter ξ_2 und ξ_3 verstehe ich die konjugierten Linearformen mit Koeffizienten aus K_2 und K_3. Die Gleichung der Fläche wird dann

$$a x^3 + \beta_1 x^2\xi_1 + \beta_2 x^2\xi_2 + \beta_3 x^2\xi_3 + x(\gamma_1\xi_2\xi_3 + \gamma_2\xi_1\xi_3 + \gamma_3\xi_1\xi_2) + \\ + d\xi_1\xi_2\xi_3 = 0,$$

wobei β_i und γ_i für $i=1, 2, 3$ konjugierte Zahlen aus K_1, K_2, K_3 sind, a und d rational. Setzt man nun

$$\xi_i + \frac{\gamma_i}{d}x = \bar{\xi}_i,$$

so geht die Gleichung über in

$$\bar{a}x^3 + (\bar{\beta}_1\bar{\xi}_1 + \bar{\beta}_2\bar{\xi}_2 + \bar{\beta}_3\bar{\xi}_3)x^2 + d\bar{\xi}_1\bar{\xi}_2\bar{\xi}_3 = 0;$$

wobei hier

$$\bar{a} = a - S\frac{\beta_1\gamma_1}{d} + \frac{2N\gamma_1}{d^2}, \quad \bar{\beta}_1 = \beta_1 - \frac{\gamma_2\gamma_3}{d}, \ldots$$

gesetzt ist. Sind die $\bar{\beta}_r$ alle $=0$, hat man eine Gleichung, deren allgemeine Lösung sofort angebbar ist. Ich sehe deshalb davon ab. Setzt man nun

$$\frac{\bar{\beta}_i\bar{\xi}_i}{x} = \eta_i, \quad \bar{d} = \frac{d}{\bar{\beta}_1\bar{\beta}_2\bar{\beta}_3},$$

so nimmt unsere Gleichung die Gestalt

$$\bar{a} + S\eta + \bar{d}N\eta = 0$$

an. Ich werde nun die Gleichung

$$N\tau + hS\tau + k = 0 \tag{9}$$

zuerst in dem Sonderfall behandeln, wo die drei konjugierten Körper K_1, K_2, K_3, welche von den Wurzeln τ herrühren, rein-kubisch sind. Danach werde ich bei beliebig gelassenem Körper K_1 den Fall betrachten, wo die Wurzeln der Gleichung

$$t^3 + 3ht + k = 0 \tag{10}$$

zu rein kubischen Körpern Anlaß geben. Zuletzt betrachte ich den allgemeinen Fall. Ich werde zeigen, wie man mindestens in den beiden zuerst erwähnten Fällen eine vollständige Parameterdarstellung der rationalen Lösungen angeben kann, sofern solche überhaupt existieren.

Es sei K_1 der Körper $k(\vartheta)$, $\vartheta^3 = m$. Soll (9) von einer Zahl $\tau \in k(\vartheta)$ erfüllt sein, so kann man τ in der Form

$$\tau = \frac{a + b\vartheta}{c + d\vartheta}$$

schreiben, und man bekommt

$$N(a + b\vartheta) + hS(a + b\vartheta)(c + \varrho d\vartheta)(c + \varrho^2 d\vartheta) + kN(c + d\vartheta) = 0$$

oder

$$a^3 + mb^3 + 3h(ac^2 + mbd^2) + k(c^3 + md^3) = 0,$$

d.h.

$$a^3 + 3hac^2 + kc^3 + m(b^3 + 3hbd^2 + kd^3) = 0. \tag{11}$$

Dies bedeutet, daß m die Norm einer Zahl des Körpers ist, der von einer Wurzel α der Gl. (10) herrührt. In der Tat bedeutet (11) dasselbe wie

$$-m = N\frac{a - c\alpha}{b - d\alpha}. \tag{12}$$

Die Gl. (9) ist also stets und nur dann lösbar, wenn (12) es ist, und das kann bekanntlich immer entschieden werden. (Näheres über die Lösbarkeit weiter unten.) Die allgemeine Lösung von

$$N\Theta = -m$$

im Körper $k(\alpha)$ ist gegeben durch

$$\Theta = \Theta_0 \Phi,$$

wo Θ_0 eine spezielle Lösung ist, während Φ die allgemeine Lösung von

$$N\Phi = 1$$

ist; diese nehme ich als bekannt an. Setzt man dann

$$\Theta = \frac{a - c\alpha}{b - d\alpha} = t_0 + t_1\alpha + t_2\alpha^2,$$

so bekommt man

$$t_2 b - t_1 d = 0, \quad a = t_0 b + k t_2 d, \quad c = t_0 d - t_1 b - 3h t_2 d,$$

d.h.

$$b = \frac{t_1}{t_2} d, \quad a = \left(\frac{t_0 t_1}{t_2} + k t_2\right) d, \quad c = \left(t_0 - \frac{t_1^2}{t_2} - 3 h t_2\right) d.$$

Folglich ist

$$\tau = \frac{t_0 t_1 + k t_2^2 + t_1 \vartheta}{t_0 t_2 - t_1^2 - 3 h t_2^2 + t_2 \vartheta}$$

die allgemeine Lösung von (9).

Danach betrachte ich den Fall, daß der Körper $k(\alpha) = k(\vartheta)$ ist, $\vartheta^3 = m$, m rational. Dann müssen rationale Zahlen a, b, c, d derart existieren, daß

$$-\vartheta = \frac{a\alpha + b}{c\alpha + d}$$

und folglich

$$(a\alpha + b)^3 + m(c\alpha + d)^3 = 0$$

ist, woraus

$$a^3 + m c^3 : a^2 b + m c^2 d : a b^2 + m c d^2 : b^3 + m d^3 = 1 : 0 : h : k$$

folgt. Wegen

$$N(a\tau + b) = a^3 N\tau + a^2 b S\tau\tau' + a b^2 S\tau + b^3$$
$$m N(c\tau + d) = m(c^3 N\tau + c^2 d S\tau\tau' + c d^2 S\tau + d^3)$$

erkennt man dann, daß (9) in der Form

$$N\Theta_1 + m N\Theta_2 = 0$$

mit

$$\Theta_1 = a\tau + b, \quad \Theta_2 = c\tau + d$$

geschrieben werden kann. Ist also Θ die allgemeine Lösung von

$$N\Theta = -m,$$

so ist

$$\tau = -\frac{b - d\Theta}{a - c\Theta}$$

die allgemeine Lösung von (9).

Im allgemeinen Falle kann man folgendermaßen vorgehen. Es sei $K_1 = k(\vartheta_1)$, ϑ_1 eine Wurzel der Gleichung

$$\vartheta^3 + c_1\vartheta^2 + c_2\vartheta + c_3 = 0, \tag{13}$$

deren Diskriminante D heißen möge. Dann ist der Körper $K_1(\sqrt{-3D}) = k(\vartheta_1, \sqrt{-3D})$ rein kubisch in bezug auf $k(\sqrt{-3D})$. Man kann deshalb in derselben Weise wie oben die allgemeine Lösung von (9) mit Koordinaten im Körper $k(\sqrt{-3D})$ finden, also die allgemeine Lösung von (9) mit τ in $K_1(\sqrt{-3D})$. Zieht man nun eine Sekante durch ein beliebiges Paar von Punkten P_1 und P_2, deren Koordinaten konjugierte Zahlen in $k(\sqrt{-3D})$ sind, so wird der dritte Schnittpunkt mit der Fläche ein rationaler Punkt P. Ob man aber in dieser Weise wirklich die allgemeine Lösung in rationalen Zahlen findet, muß hier offen gelassen werden. Ich habe leider keine Zeit gehabt, auf diese Frage ernstlich einzugehen.

SEGRE [5] gibt an, wie man eine vollständige Parameterdarstellung der rationalen Punkte erhalten kann, wenn man zuerst zwei rationale Punkte P und Q gefunden hat. Sind O_1, O_2, O_3 die drei singulären Punkte, so sieht man leicht ein, daß von den fünf Punkten O_1, O_2, O_3, P, Q nie vier in einer Ebene liegen. Jede windschiefe kubische Kurve durch O_1, O_2, O_3, P, Q mit rationalen Koeffizienten schneidet dann die kubische Fläche in einen einzigen weiteren, also rationalen, Punkt. Umgekehrt geht durch jeden rationalen Punkt der Fläche eine solche kubische Kurve.

2. Mittels linearer Transformation mit rationalen Koeffizienten kann man erreichen, daß die Doppelpunkte die Koordinaten $0, 0, \sqrt{m}$ und $0, 0, -\sqrt{m}$ bekommen, m rational. Die ganze z-Achse liegt dann in der Fläche. Legt man eine Ebene $y = kx$ durch die z-Achse, k rational, so schneidet diese die kubische Fläche in einen Kegelschnitt. In bekannter Weise kann man dann für jeden Wert von k entscheiden, ob rationale Punkte auf dem Kegelschnitt existieren, und auch eine allgemeine Formel für diese aufstellen. Man bekommt jedoch hierdurch nicht eine Parameterdarstellung der rationalen Lösungen überhaupt, weil man gewöhnlich nicht eine solche Darstellung für die Werte derjenigen k hat, für welche Lösbarkeit besteht. Es ist aber klar, daß man genau wie im Fall 1 hier eine vollständige Parameterdarstellung der Lösungen im Körper $k(\sqrt{m})$ aufstellen kann. Zieht man dann die Sekante zwischen zwei beliebigen konjugierten Punkten, so wird der dritte Schnittpunkt rational. Allerdings bleibt es hier wiederum fraglich, ob man alle rationalen Punkte auf der Fläche in dieser Weise erhält.

Gibt es einen dritten oder sogar einen vierten Doppelpunkt, so kann man leicht eine vollständige Parameterdarstellung erhalten. Gibt es genau drei Doppelpunkte, so ist ja der dritte rational, und man gelangt zum Fall 1 zurück. Gibt es zwei weitere Doppelpunkte, so sind diese entweder rational, so daß man wieder den Fall 1 hat, oder sie sind auch konjugiert quadratisch irrational. In dem letzten Fall kann man aber so vorgehen wie im Fall 4.

Es geht aus dem obigen hervor, daß in den Fällen 1, 2, 4 stets rationale Punkte auf der Fläche vorkommen. Im Fall 3 ist es aber nicht immer so. MORDELL ([3], S. 6) hat die Vermutung aufgestellt, daß die Gleichung $f(x, y, z, u) = 0$, wo f eine kubische Form ist, dann — und selbstverständlich auch nur dann — eine rationale Lösung hat, wenn die Kongruenz $f(x, y, z, u) \equiv 0$ für jeden Modul lösbar ist. *Ich werde hier zeigen, daß diese Vermutung richtig ist im Fall 3 und also überhaupt für kubische Flächen, die mindestens einen singulären Punkt besitzen.* (E. S. SELMER [6] hat die Richtigkeit der Vermutung für die Flächen $a_1 x_1^3 + a_2 x_2^3 + a_3 x_3^3 + a_4 x_4^3 = 0$ unter der Voraussetzung bewiesen, daß $a_3 a_4 / a_1 a_2$ ein rationaler Kubus ist.)

Ich stütze den Beweis auf die folgende bekannte Tatsache. Ist K ein zyklischer kubischer Oberkörper über dem Körper k und ist $m \in k$, so besitzt die Gleichung

$$N\tau = m \tag{14}$$

eine Lösung $\tau \in K$, wenn die entsprechende Kongruenz

$$N\tau \equiv m \tag{15}$$

für jeden Modul lösbar ist (vgl. HASSE [2], S. 40). Es sei jetzt K ein nichtzyklischer kubischer Oberkörper über k. Dann ist die normale Erweiterung $\bar{K}$ vom sechsten Grade über k zyklisch über einen relativ quadratischen Oberkörper k_2 über k. Ist nun (15) lösbar für jeden Modul im Körper k, so ist dies um so mehr der Fall im Körper k_2. Nach dem zitierten Satz muß dann (14) in $\bar{K}$ lösbar sein. Dies bedeutet, daß die kubische Gleichung für die Koeffizienten von τ, wenn τ als Linearform mittels einer Basis in k ausgedrückt ist, in Zahlen aus k_2 lösbar ist. Die Anwendung der Sekantenmethode zeigt dann, daß sie auch in k lösbar sein muß. Aus dieser Verallgemeinerung folgert man leicht die Gültigkeit der MORDELLschen Vermutung in den zwei Sonderfällen von 3, die oben erwähnt sind. Im allgemeinen Falle kann man so schließen: Es sei wie früher D die Diskriminante von (13). Es werde angenommen, daß die Kongruenz

$$N\tau + hS\tau + k \equiv 0 \tag{16}$$

im absolut rationalen Zahlkörper für jeden Modul lösbar ist. Dann ist sie um so mehr lösbar in $k(\sqrt{-3D})$ für jeden Modul. Im Körper $k(\sqrt{-3D})$ ist aber (9) mit (12) gleichwertig, wo jetzt m durch eine Zahl μ in $k(\sqrt{-3D})$ ersetzt ist. Die Kongruenz (16) ist mit einer entsprechenden Kongruenz

$$N\Theta_1 + \mu N\Theta_2 \equiv 0$$

gleichwertig. Da also diese für jeden Modul lösbar ist, so ist (12) lösbar in $k(\sqrt{-3D})$, woraus folgt, daß (9) in demselben Körper lösbar ist. Mittels der Sekantenmethode sieht man dann ein, daß (9) auch rational lösbar ist.

§ 4.

Ich werde nun die allgemeinen kubischen Flächen betrachten, aber mit Ausschluß der Regelflächen.

Es verdient bemerkt zu werden, daß die Unterscheidung zwischen den Flächen ohne singuläre Punkte und den Flächen mit solchen für das Problem der rationalen Punkte nicht immer wesentlich ist. *In der Tat kommt es oft vor, daß eine birationale Transformation mit rationalen Koeffizienten existiert, die eine Fläche ohne singuläre Punkte in eine mit solchen überführt.* Ich gebe ein Beispiel.

Die kubische Fläche

$$x^3 + 2y^3 + 3z^3 + 6u^3 = 0 \tag{17}$$

besitzt keinen singulären Punkt. Aber durch die Transformation

$$x : y : z : u = -t_0 t_1 + 2t_2^2 : -t_1^2 + t_0 t_2 : -t_1 t_3 : t_2 t_3$$

geht sie über in

$$t_0^3 + 2t_1^3 + 4t_2^3 - 6t_0 t_1 t_2 + 3t_3^3 = 0, \tag{18}$$

und diese Fläche hat die drei Doppelpunkte

$$t_0 : t_1 : t_2 : t_3 = \begin{cases} \vartheta^2 : \vartheta : 1 : 0 \\ \varrho^2\vartheta^2 : \varrho\vartheta : 1 : 0 \\ \varrho\vartheta^2 : \varrho^2\vartheta : 1 : 0 \end{cases}$$

mit $\vartheta^3 = 2$, $\varrho^2 + \varrho + 1 = 0$. Die umgekehrte Transformation ist

$$t_0 : t_1 : t_2 : t_3 = xz^2 + 2yu^2 : yz^2 - xzu : xu^2 - yzu : z^3 + 2u^3.$$

Da die allgemeine Lösung von (18) leicht angebbar ist, so findet man hierdurch leicht die allgemeine Lösung von (17).

Man könnte fragen, ob nicht vielleicht jede kubische Fläche birational in eine solche mit singulären Punkten transformierbar ist. Wie nun dies sein mag, so wird es wohl ein Problem sein, die betreffenden Transformationen zu finden; deshalb dürfte es wohl auf alle Fälle berechtigt sein, die Flächen zu studieren, die keinen Doppelpunkt besitzen.

Bekanntlich kann man in dem Falle, wo die Fläche zwei rationale Geraden l_1 und l_2 enthält, die nicht in derselben Ebene liegen, in trivialer Weise eine vollständige Parameterdarstellung ihrer rationalen Punkte aufstellen. In der Tat, eine Gerade, die einen rationalen Punkt auf l_1 mit einem solchen auf l_2 verbindet, schneidet die Fläche in einen dritten rationalen Punkt. Ist umgekehrt P ein rationaler Punkt der Fläche, so liegt P zugleich mit l_1 in einer rationalen Ebene, welche l_2 in einen rationalen Punkt P_2 schneidet. Die Gerade welche P_2 mit P verbindet, schneidet dann l_1 in einen rationalen Punkt P_1.

Abgesehen von diesem Falle aber scheint es meistens schwierig, eine vollständige Darstellung der rationalen Punkte zu geben. Man hat sich dann damit begnügt, Scharen von Lösungen mit einem oder mehreren Parametern zu finden, ohne auf die Frage einzugehen, ob sie alle Lösungen enthalten. In erster Instanz hat man nun beweisen können, daß, falls ein rationaler Punkt auf der Fläche vorkommt, unendlich viele solche vorhanden sind. Für diesen Satz haben SEGRE und R. F. WHITEHEAD Beweise gegeben. Der Beweis WHITEHEADs wird l. c. von MORDELL wiedergegeben mit der Bemerkung, daß er so elementar ist, daß es überraschend ist, daß er nicht früher entdeckt worden ist, z. B. von EULER. Es ist richtig, daß der Beweis elementar ist, aber er beruht doch auf einem Kunstgriff. Ich möchte deshalb einen anderen Beweis vorschlagen, der mir natürlicher zu sein scheint.

Bekanntlich wird die Tangentenebene T in einem Punkt P gewöhnlich die Fläche in einer Kurve C dritten Grades mit Doppelpunkt in P schneiden. Ist P rational, so ist T eine rationale Ebene, und jede rationale Gerade durch P in T schneidet dann C in einen rationalen Punkt. Dies gibt sofort eine einparametrige Schar von rationalen Punkten. Nun kann es sein, daß C zerfällt in eine Gerade und einen Kegelschnitt durch P; diese sind dann rational, so daß wir wiederum unendlich viele rationale Punkte erhalten. Es kann aber auch eintreten, daß C sich zerlegt in drei Geraden. Ist mindestens eine dieser Geraden rational, so gibt es wieder unendlich viele rationale

Punkte. Es bleibt deshalb nur der Fall, daß T die Fläche in drei Geraden l_1, l_2, l_3 schneidet, welche konjugiert kubisch irrational sind. Ferner müssen l_1, l_2, l_3 alle durch P gehen; denn sonst wären die drei Schnittpunkte dieser Geraden alle rational, da sie konjugiert sind, und dann wären alle drei Geraden rational. Es kann dann auch sofort bemerkt werden, daß jede sonstige Ebene durch P die Fläche in einer Kurve schneiden muß, die in P einen Wendepunkt hat. Ich nenne einen Punkt P einen Ausnahmepunkt, wenn die zugehörige Tangentenebene T die Fläche in drei Geraden l_1, l_2, l_3 schneidet, die konjugiert in kubischen Körpern sind. Es ist klar, daß nur endlich viele Ausnahmepunkte existieren können; denn es gibt ja auf der Fläche höchstens 27 Geraden.

Es sei also P ein rationaler Ausnahmepunkt. Ich lege eine rationale Gerade l durch P. Diese schneidet die kubische Fläche in zwei Punkten P_1 und P_2, deren Koordinaten konjugiert in einem quadratischen Körper K sind. Die Gerade l kann natürlich so gewählt werden, daß weder P_1 nach P_2 einer der endlich vielen Ausnahmepunkte wird. Sollten P_1 und P_2 beide rational ausfallen, so bewirkt das nur, daß in der folgenden Überlegung K einen beliebigen quadratischen Körper bedeuten kann. Wir nehmen dann eine Ebene E_1 durch l mit Koeffizienten aus K, jedoch nicht alle rational. Dann schneidet E_1 die zu P_1 gehörige Tangentialebene T_1 in einer Geraden l_1, welche die Fläche in einem Punkt Q_1 mit Koordinaten in K schneidet. Die Ebene E_2 durch l hat dann Koeffizienten in K, konjugiert zu denen der Ebene E_1; und E_2 schneidet die zu P_2 gehörige Tangentialebene T_2 in einer Geraden l_2, welche die Fläche in einen Punkt Q_2 schneidet, dessen Koordinaten zu K gehören und konjugiert zu denjenigen von Q_1 sind. Die Gerade l_3 durch Q_1 und Q_2 schneidet dann die Fläche in einem rationalen Punkt Q. Dieser ist verschieden von P. Fielen nämlich P und Q zusammen, so lägen die Geraden l und l_3 in derselben Ebene, die dann sowohl mit E_1 wie E_2 identisch sein müßte; dies ist unmöglich, da ja E_1 und E_2 verschieden sind.

Andererseits ist klar: Wenn E_1 mit rationalen Koeffizienten gewählt wird, so sind E_1 und E_2 identisch ($= E$), und dann fallen P und Q zusammen; denn in der Ebene E wird die Schnittkurve mit der Fläche, wie oben bemerkt, in P einen Wendepunkt haben, und man erkennt dann aus der bekannten Parameterdarstellung der Schnittkurve mittels elliptischer Funktionen, daß Q mit P zusammenfallen muß. Nun können wir die Ebenen E_1 und E_2 so wenig verschieden wählen als wir wünschen, indem wir nämlich ihre Koeffizienten in K beliebig nahe an rationale Zahlen wählen. Dann muß der Punkt Q beliebig nahe an P heranrücken. Dies liefert die Existenz unendlich vieler rationaler Punkte auf der kubischen Fläche.

Das in diesem Beweis angewandte Verfahren habe ich auch benutzt zur Herleitung von Polynomlösungen in r von Gleichungen der Form $f(x, y, z) = r$ [vgl. die Bemerkung über (2) in § 2]. Zuerst kann man für eine kubische Form $f(x, y, z, u)$ mit rationalen Koeffizienten den Satz aufstellen:

Sind a_1, b_1, c_1, d_1 rationale Zahlen derart, daß $f(a_1, b_1, c_1, d_1) = 0$, während (a_2, b_2, c_2, d_2) ein Punkt auf der Schnittkurve ist zwischen der Fläche $f(x, y, z, u) = 0$ und ihrer Tangentenebene in (a_1, b_1, c_1, d_1), so ist

$$f(a_1 r + a_2,\ b_1 r + b_2,\ c_1 r + c_2,\ d_1 r + d_2) = h r,$$

wo h eine rationale Zahl ist.

Die Richtigkeit dieser Behauptung sieht man sofort wegen der TAYLOR-Entwicklung

$$f(a_1 r + a_2,\ b_1 r + b_2,\ c_1 r + c_2,\ d_1 r + d_2) = f(a_1, b_1, c_1, d_1)\, r^3 +$$
$$+ \left(\frac{\partial f}{\partial x}(a_1, b_1, c_1, d_1)\, a_2 + \frac{\partial f}{\partial y}(a_1, b_1, c_1, d_1)\, b_2 + \frac{\partial f}{\partial z}(a_1, b_1, c_1, d_1)\, c_2 + \right.$$
$$\left. + \frac{\partial f}{\partial u}(a_1, b_1, c_1, d_1)\, d_2\right) r^2 + \left(\frac{\partial^2 f}{\partial x^2}(a_1, b_1, c_1, d_1)\, \frac{a_2^2}{2} + \cdots\right) r + f(a_2, b_2, c_2, d_2).$$

Hier verschwinden nämlich das erste, zweite und vierte Glied.

Als ein Beispiel sei die Form

$$x^3 + y^3 + 2z^3 + (x^2 + y^2 + 2z^2)\, u$$

gewählt. Hier können wir z. B.

$$a_1 = 1, \quad b_1 = 1, \quad c_1 = -1, \quad d_1 = 0$$

setzen. Die Tangentialebene in diesem Punkte hat die Gleichung

$$3x + 3y + 6z + 4u = 0.$$

Wählt man die Tangente, für welche

$$u = x - y$$

ist, so findet man den Schnittpunkt

$$a_2 = 5, \quad b_2 = 167, \quad c_2 = 22, \quad d_2 = -162.$$

Man erhält also hier die Identität

$$(r + 5)^3 + (r + 167)^3 + 2(-r + 22)^3 -$$
$$- ((r + 5)^2 + (r + 167)^2 + 2(-r + 22)^2)\, 162 = 2 \cdot 3^9 r,$$

oder mit anderen Worten: Man findet, daß

$$x = \frac{r + 5}{162}, \quad y = \frac{r + 167}{162}, \quad z = \frac{-r + 22}{162}$$

eine Lösung der inhomogenen Gleichung

$$x^3 + y^3 + 2z^3 - x^2 - y^2 - 2z^2 = \frac{1}{108} r$$

ist.

Von dieser Lösung aus können wir dann durch Tangenten- oder Sekantenziehung weitere Polynomlösungen ableiten, auch wenn man will mit mehr Parametern $s_1, s_2, \ldots$. Es hat kaum einen Zweck, darauf näher einzugehen.

Ich will mich hier damit begnügen, einige Ergebnisse mitzuteilen, die ich beim Studium der Gleichungen

(19) $$xy(x+y)+z^3=ru^3$$

(20) $$x^3+y^2+z^2=r$$

(21) $$x^3+y^3+z^2=r$$

(22) $$x^3+y^3+z^3=ru^3$$

erhalten habe. Für (19) habe ich sowohl die Lösung

$$x=12^3r,\quad y=-r^4+142r^2-864r+1331,\quad z=-12r^3+636r,$$
$$u=-60r^2+1452$$

wie auch

$$x=18^3r,\quad y=(r+1)(r^2-34r-1331),\quad z=-108r(r-35),$$
$$u=18(r^2-94r+121)$$

gefunden. Eine Lösung von (20) ist

$$x=1/4(-r^2+2r-2),\quad y=1/8(-r^2+6r+2),$$
$$z=1/8(-r^3+3r^2-4r-2).$$

Lösungen von (21) sind

$$x=-1/27(r+27),\quad y=1/27(r-54),\quad z=1/9(r+27)$$

und

$$x=-1/63(27r-26),\quad y=1/63(27r-47),\quad z=1/63(27r+37).$$

Was (22) betrifft, möchte ich etwas näher darauf eingehen, wie ich eine Lösung erhalten habe mit Hilfe des Verfahrens, das ich oben benutzt habe, um unendlich viele rationale Lösungen abzuleiten; ausgehend von einem Punkt, dessen zugehörige Tangentialebene die Fläche in drei kubisch irrationalen konjugierten Geraden schneidet. Im Falle der Gl. (22) ist P mit den Koordinaten $-1, 0, 1, 0$ ein solcher Punkt. Die Gerade

$$2y=x+z$$

geht durch P und schneidet die Fläche (22) in den beiden folgenden Punkten P_1 und P_2:

$$P_1:\quad x=1/3(1+2\sqrt{-2}),\quad y=1/3(2+\sqrt{-2}),\quad z=1,\quad u=0$$
$$P_2:\quad x=1/3(1-2\sqrt{-2}),\quad y=1/3(2-\sqrt{-2}),\quad z=1,\quad u=0.$$

Die Tangentenebene in P_1 ist

$$1/9(-7+4\sqrt{-2})x+1/9(2+4\sqrt{-2})y+z=0.$$

Ich wähle für l_1 die Schnittgerade dieser Ebene und der Ebene

$$4u=\sqrt{-2}(x-2y+z).$$

Man findet dann nach einiger Rechnung, daß l_1 die Fläche in den Punkt Q_1 mit den Koordinaten $x:y:z:u=$

$$(1+8\sqrt{-2}+r(4+2\sqrt{-2})):(-10\sqrt{-2}+4r):$$
$$:(-1+8\sqrt{-2}+r(4-2\sqrt{-2})):-18.$$

schneidet. Die Gerade zwischen Q_1 und dem konjugierten Punkt Q_2 kann dann dargestellt werden durch die Gleichung

$$x:y:z:u=(1+4t+r(4+t)):(-5t+4r):(-1+4t+r(4-t)):-18.$$

Also bekommt man für die Schnittpunkte mit der Fläche die Gleichung

$$(1+4t+r(4+t))^3+(-5t+4r)^3+(-1+4t+r(4-t))^3+18^3r=0,$$

wobei zwei der Wurzeln $2\sqrt{-2}$ und $-2\sqrt{-2}$ sind. Man findet, daß die dritte Wurzel $(-24r^3-732r)/(24r+3)$ ist. Die Einsetzung hiervon liefert den Punkt Q mit Koordinaten

$$x:y:z:u=(-8r^4-236r^2-972r+1):(72r^3+1224r):$$
$$:(8r^4+236r^2-972r-1):-18(8r^2+1),$$

wodurch eine Polynomlösung von (22) gefunden ist.

Bekanntlich hat RICHMOND auf Grund einer früheren Arbeit von RYLEY bewiesen, daß für $r>0$ positive rationale Lösungen x, y, z der Gleichung

$$x^3+y^3+z^3=r$$

existieren. Seine Methode besteht in rechnerischen Kunstgriffen (vgl. HARDY and WRIGHT [1], S. 196—198). Es bietet keine Schwierigkeit, mittels wiederholter Tangentenziehung positive rationale Lösungen der oben erwähnten Gleichungen zu finden, wenn man überhaupt rationale Lösungen gefunden hat; ich will aber hier nicht näher darauf eingehen.

Literatur.

[1] HARDY, G. H., and E. M. WRIGHT: An Introduction to the Theory of Numbers 1938. — [2] HASSE, H.: Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper. Teil II: Reziprozitätsgesetz. Jber. dtsch. Math.-Ver., Erg.-Bd. **6** (1930). — [3] MORDELL, L. J.: Rational points on cubic surfaces. Publ. Math. Debrecen **1**, 1—6 (1949). — [4] SEGRE, B.: A parametric solution of the indeterminate cubic equation $z^2=f(x, y)$. J. London Math. Soc. **18**, 226—233 (1943). — [5] SEGRE, B.: On arithmetical properties of singular cubic surfaces. J. London Math. Soc. **19**, 84—91 (1944). — [6] SELMER, E. S.: Sufficient congruence conditions for the existence of rational points on certain cubic surfaces. Math. Scand. **1**, 113—119 (1953). — [7] SKOLEM, TH.: Diophantische Gleichungen. Erg. Math. **5**, H. 4 (1938). — [8] SKOLEM, TH.: On the existence of a multiplicative basis for an arbitrary algebraic field. Det Kong. Norske Vidensk. Selsk., Forh. **20**, Nr. 2 (1947).

Guldbergs vei 4, Smestad, Oslo (Norwegen).

(Eingegangen am 16. April 1955.)

HAHN, W.
Math. Zeitschr. Bd. 63, S. 313—319 (1955)

Über analytische Lösungen linearer Differential-Differenzengleichungen.

ISSAI SCHUR zum Gedächtnis.

Von
WOLFGANG HAHN.

1. Wir betrachten im folgenden die Lösungen einer linearen Differential-Differenzengleichung

$$\sum_{i=0}^{p} a_i Y^{(i)}(t) + \sum_{i=0}^{q} b_i Y^{(i)}(t-1) = 0 \qquad (p \geq q) \tag{1}$$

mit komplexen Koeffizienten a_i, b_i. Da das Studium derartiger Gleichungen vornehmlich durch technische Fragestellungen angeregt worden ist, in denen die Zeit als unabhängige Veränderliche t auftritt, hat man bisher fast ausschließlich die Lösungen als Funktionen einer reellen Veränderlichen untersucht. Analytische Lösungen von (1) sind bisher nur von Herrn LEONTEV[1]) behandelt worden. Die nachstehende Arbeit beschäftigt sich ebenfalls mit analytischen Lösungen, und zwar in dem Sonderfall $p > q$, in dem die Verhältnisse viel einfacher sind als in dem allgemeinen Fall $p \geq q$. Die Ergebnisse liegen aber in anderer Richtung als die LEONTEVschen [bis auf die Formel (9), die auch bei LEONTEV steht]; es handelt sich vor allem um das Verhalten gewisser unendlicher Reihen auf der Grenze ihres Konvergenzbereichs, um die Übertragung der klassischen Sätze von Abel und Tauber und um den Nachweis der Existenz singulärer Punkte.

Mit $C_1, C_2, \ldots$ sind im folgenden positive Konstanten bezeichnet, auf deren genauen Wert es nicht weiter ankommt.

2. Aus der Theorie der Gl. (1) im Reellen benötigen wir die folgenden Ergebnisse[2]).

a) Die Gleichung hat unendlich viele Partiallösungen der Gestalt

$$Y_j(t) = e^{r_j t} \qquad (j = 1, 2, \ldots). \tag{2}$$

Dabei sind die Zahlen r_j die Nullstellen der „charakteristischen Funktion"

$$A(r) = \sum_{i=0}^{p} a_i r^i + e^{-r} \sum_{i=0}^{q} b_i r^i. \tag{3}$$

[1]) LEONTEV, A. F.: Differential-Differenzengleichungen. Mat. Sbornik **24** (66), 347 bis 374 (1949) [Russisch]. Eine kurze Zusammenstellung der Hauptergebnisse steht bei HAHN, W.: Bericht über Differential-Differenzengleichungen mit festen und veränderlichen Spannen. Jber. dtsch. Math.-Ver. **57**, 55—84 (1954).

[2]) WRIGHT, E. M.: The linear difference-differential equation with constant coefficients. Proc. Roy. Soc. Edinburgh A **62**, 387—393 (1949).

Wenn r_k m-fache Nullstelle von $A(r)$ ist, sind auch die Funktionen

$$t e^{r_k t},\ t^2 e^{r_k t},\ \ldots,\ t^{m-1} e^{r_k t} \tag{4}$$

Lösungen von (1). Es können aber höchstens endlich viele Mehrfachnullstellen auftreten.

b) Jede für reelle positive t erklärte Lösung von (1) läßt sich als lineare Verbindung der Partiallösungen (2) bzw. (4) darstellen, also in der Gestalt

$$f(t) = \sum_{j=1}^{\infty} A_j(t)\, e^{r_j t}. \tag{5}$$

Die Ausdrücke $A_j(t)$ sind im allgemeinen Konstanten und nur für endlich viele Indizes, die den Mehrfachnullstellen entsprechen, Polynome.

c) Ordnet man die Nullstellen von $A(r)$, die positiven Imaginärteil haben — sie sollen nachstehend durch $s_n = u_n + i v_n$ bezeichnet werden — nach wachsenden Beträgen, so gilt für große n eine asymptotische Darstellung[3])

$$s_n = -(p-q)\log n + C_1 + C_2 n i + C_3 i + R_n. \tag{6}$$

Das Restglied R_n genügt einer Abschätzung

$$|R_n| < C_4 \frac{\log n}{n}. \tag{7}$$

Die asymptotische Darstellung (6) gilt sinngemäß auch für die Nullstellen mit negativem Imaginärteil.

3. Da sich jede Lösung der Funktionalgleichung, die für $t > 0$ erklärt ist, als Reihe der Form (5) darstellen läßt, gilt dies auch für jede analytische Lösung; denn man kann durch eine lineare Transformation der Variablen, die die Form der Reihe (5) ungeändert läßt, stets erreichen, daß die Funktion in einem Intervall der positiven reellen Achse regulär ist. Die Partiallösungen (2) und (4) sind natürlich ganze Funktionen. Wir betrachten deshalb nur unendliche Reihen und beschränken uns zunächst auf Nullstellen s_n mit positivem Imaginärteil v_n:

$$f(z) = \sum_{n=1}^{\infty} A_n e^{s_n z}. \tag{8}$$

Die unabhängige Variable, die jetzt komplex ist, ist mit z bezeichnet, die A_n sind beliebige komplexe Zahlen. Von den endlich vielen Mehrfachnullstellen ist abgesehen, da die ihnen entsprechenden Partiallösungen (4) auf das analytische Verhalten der Funktion keinen Einfluß haben. Es wird sich zeigen, daß sich die Reihen (8) in mancher Hinsicht wie Dirichletsche Reihen verhalten. Sie konvergieren, wenn sie nicht eine ganze Funktion darstellen, in einer Halbebene, die durch eine Konvergenzhorizontale nach unten begrenzt wird. Bei Annäherung an Punkte dieser Geraden gilt ein Satz vom Abelschen Typ — aus dem man folgern kann, daß die Reihe in der offenen

[3]) LANGER, R. E.: The asymptotic location of the roots of a certain transcendental equation. Trans. Amer. Math. Soc. **31**, 837—844 (1929).

Halbebene eine analytische Funktion darstellt — und unter Zusatzannahmen über die Koeffizienten ein Satz vom Tauberschen Typ. Mit seiner Hilfe kann man unter gewissen Voraussetzungen über die Koeffizienten die Existenz singulärer Punkte auf der Konvergenzgeraden zeigen.

4. Bestimmung des Konvergenzgebiets. Es sei

$$a = \limsup_{n\to\infty} \frac{\log|A_n|}{v_n} \qquad (s_n = u_n + i v_n). \tag{9}$$

Dann konvergiert die Reihe (8) *für* $Im\, z > a$ *absolut und divergiert für* $Im\, z < a$.

Der *Beweis* verläuft wie der des entsprechenden Satzes über Dirichlet-Reihen. Gibt man $\varepsilon > 0$ beliebig klein vor, so ist nach (9) zuletzt $\log|A_n| \leq (a+\varepsilon) v_n$ und unendlich oft $\log|A_n| > (a-\varepsilon) v_n$. Nun sei $z = x + iy$. Dann ist für hinreichend große n

$$|A_n e^{s_n z}| \leq e^{a v_n} e^{\varepsilon v_n} e^{u_n x - v_n y}.$$

Wählt man $y > a$ und $0 < \varepsilon < y - a$, also etwa $y = a + \varepsilon + b$ mit $b > 0$, so ist

$$|A_n e^{s_n z}| \leq e^{u_n x - b v_n}.$$

Aus (6) ergibt sich, daß die rechte Seite dieser Ungleichung in der Form

$$C_5 n^{-(p-q)x} e^{-C_3 b n} (1 + R_n')$$

geschrieben werden kann; für R_n' gilt eine Abschätzung wie (7). Da $C_2 b > 0$ ist, ist die Reihe mit dem allgemeinen Glied $A_n e^{s_n z}$ absolut konvergent. Ist aber $y < a$, etwa $y = a - 2\varepsilon$, $\varepsilon > 0$, so ist unendlich oft

$$|A_n e^{s_n (x+iy)}| > e^{(a-\varepsilon) v_n + u_n x + 2\varepsilon v_n - a v_n} = e^{u_n x + \varepsilon v_n},$$

und das strebt gegen unendlich.

Über das Verhalten der Reihe auf der Konvergenzhorizontalen (KH) gilt: *Es sei*

$$c = \limsup_{n\to\infty} \frac{\log|A_n| - a v_n}{-u_n}. \tag{10}$$

Dann konvergiert die Reihe $f(z)$ *auf der Halbgeraden* $z = x + ia$, $x > c + \frac{1}{p-q}$, *absolut und divergiert für* $x < c$.

Beweis. Man erkennt wie eben, daß zuletzt für jedes $c' > c$

$$|A_n e^{s_n (x+ia)}| = |A_n|\,|e^{u_n x - a v_n}| \leq e^{u_n (x-c')}$$

ist. Wegen (6) ist

$$e^{u_n (x-c')} = C_6 n^{-(p-q)(x-c')} (1 + R_n'');$$

die Reihe konvergiert daher absolut, wenn $(x-c')(p-q) > 1$ bzw., da das für alle $c' > c$ gelten soll, wenn $(x-c)(p-q) > 1$ ist. Wenn andererseits $f(z)$ für den Punkt $z_0 = x_0 + i y_0$ der KH konvergiert, so ist sicher $A_n e^{s_n z_0}$ beschränkt, $|A_n e^{s_n z_0}| < M$. Dann wird

$$|f(z)| = \left|\sum_{n=1}^{\infty} A_n e^{s_n z_0} e^{s_n (z-z_0)}\right| \leq M \sum_{n=1}^{\infty} |e^{s_n (z-z_0)}|.$$

Wenn $Im(z-z_0)>0$ ist, konvergiert die rechtsstehende Reihe; das liefert nichts Neues. Ist $Im(z-z_0)=0$, d.h. liegt z auf der KH, so konvergiert die Reihe im Fall $Re(z-z_0)>\frac{1}{p-q}$ absolut. Das Intervall, in dem Punkte bedingter Konvergenz auftreten, kann also höchstens die Länge $\frac{1}{p-q}$ haben.

Die durch (9) und (10) definierten Zahlen können auch unendlich große Werte annehmen, wie aus der folgenden Zusammenstellung hervorgeht.

A_n	a	c
e^{n^2}	$+\infty$	$-$
e^{-n^2}	$-\infty$	$-$
$e^{-\sqrt{n}}$	0	$-\infty$
1	0	0
$n^{\log n}$	0	$+\infty$

An der Stelle $z=c$ kann Konvergenz oder Divergenz vorliegen, wie die Beispiele $A_n=1$ und $A_n=\frac{(-1)^n}{\log n}$ zeigen. In beiden Fällen ist $c=0$; während die erste Reihe divergiert, konvergiert die zweite nach dem Leibnizschen Kriterium.

5. Satz vom Abelschen Typus. *Es sei $f(z)$ an der Stelle $z=z_0$ der KH konvergent. Dann konvergiert die Reihe gleichmäßig in jedem Winkelraum mit dem Scheitelpunkt z_0, dessen Öffnung kleiner als π ist und der ganz der oberen Halbebene angehört.*

Für den Beweis führen wir die folgenden Abkürzungen ein:

$$z-z_0=\zeta=\xi+i\eta, \qquad b_k=e^{s_k\zeta},$$

$$a_{m,k}=\sum_{j=m}^{k}A_j e^{s_j z_0}, \quad \text{also} \quad A_k e^{s_k z_0}=a_{m,k}-a_{m,k-1} \qquad (k>m),$$

$$\sigma_{m,M}=\sum_{k=m+1}^{m+M}A_k e^{s_k z}=\sum_{k=m+1}^{m+M}A_k e^{s_k z_0} b_k.$$

Durch partielle Summation erhält man

$$\sigma_{m,M}=a_{m,m+M}\,b_{m+M}-\sum_{t=m+1}^{m+M-1}a_{m,t}(b_{t+1}-b_t).$$

Da $f(z_0)$ konvergiert, gehen die Teilsummen $a_{m,k}$ mit $m\to\infty$ gleichmäßig bezüglich k gegen Null, d.h. es ist

$$|a_{m,m+M}|<B_m \quad \text{mit} \quad \lim B_m=0 \quad \text{für} \quad m\to\infty.$$

Daher wird

$$|\sigma_{m,M}|\leq B_m|b_{m+M}|+B_m\sum_{t=m+1}^{m+M-1}|b_{t+1}-b_t|,$$

und der Nachweis der gleichmäßigen Konvergenz der Reihe $f(z)$ ist erbracht, wenn gezeigt worden ist, daß die Reihe

$$\sum_{t=m+1}^{\infty} |b_{t+1} - b_t| \tag{11}$$

innerhalb des Winkelraums gleichmäßig konvergiert. Es ist

$$b_{t+1} - b_t = e^{s_t \zeta}(e^{(s_{t+1}-s_t)\zeta} - 1). \tag{12}$$

Die Klammer schätzen wir nach der Formel

$$|1 - e^w| = \left| \int_0^w e^x \, dx \right| \leq |w|$$

ab, die im Fall $Re\, w < 0$ gilt und hier angewandt werden darf, da

$$Re\,[(s_{t+1} - s_t)\zeta] = (u_{t+1} - u_t)\,\xi - (v_{t+1} - v_t)\,\eta$$

mit Rücksicht auf (6) für große t sicher negativ ist. Ferner ist

$$|e^{s_t \zeta}| = e^{u_t \xi - v_t \eta}.$$

Im Winkelraum ist

$$\eta \geq C_7 |\xi|,$$

und wieder wegen (6) ist für große t

$$u_t \xi - v_t \eta < - C_8\, t\, \eta,$$

also

$$|e^{s_t \zeta}| < e^{-C_8 t \eta}.$$

Setzt man das alles in (12) ein, so folgt

$$|b_{t+1} - b_t| < |\zeta|\,|s_{t+1} - s_t|\, e^{-C_8 t \eta}.$$

Ferner ist im Winkelraum $|\zeta| \leq C_9 \eta$. Beachtet man noch, daß $s_{t+1} - s_t$ beschränkt ist, so folgt schließlich

$$|b_{t+1} - b_t| < C_{10}\, \eta\, e^{-C_8 t \eta}.$$

Summiert man nun über alle hinreichend großen t, so erhält man für die Reihe (11) eine Abschätzung

$$\sum_{t=m+1}^{\infty} |b_{t+1} - b_t| < C_{11} \frac{\eta}{1 - e^{-C_8 \eta}}.$$

Die rechte Seite der Ungleichung hat für $\eta \to 0+$, was der Annäherung von z gegen z_0 innerhalb des Winkelraums entspricht, einen endlichen Grenzwert. Damit ist der Beweis erbracht.

Aus dem Satz ergibt sich, daß $f(z)$ bei Annäherung an einen Konvergenzpunkt stetig ist, sofern die Annäherung im Winkelraum stattfindet. Es ergibt sich weiter, daß die Reihe $f(z)$ in jeder Halbebene $Im\, z \geq a' > a$ eine analytische Funktion darstellt.

Wenn z_0 ein Punkt absoluter Konvergenz ist, kann man noch die rechts von z_0 gelegene Hälfte der KH in den Winkelraum mit einbeziehen. Es ist ja dann

$$|f(z)| = \left|\sum_{n=1}^{\infty} A_n e^{s_n z_0} e^{s_n \zeta}\right| \leq \sum_{n=1}^{\infty} |A_n e^{s_n z_0}|\,|e^{s_n \zeta}|.$$

Im Winkelraum ist $e^{s_n \zeta}$ beschränkt, wenn ζ beschränkt bleibt; man kann also $f(z)$ durch eine absolut konvergente Reihe majorisieren.

6. Satz vom Tauberschen Typ. *Es sei die durch* (12) *erklärte Zahl* $a=0$, *also* $f(z)=\sum_{n=1}^{\infty} A_n e^{s_n z}$ *für* $Im\, z>0$ *konvergent; es sei* $\lim_{n\to\infty} n A_n = 0$. *Es sei* $\lim f(x_0+iy)$ *für* $y\to 0+$, *also für Annäherung von oben her, vorhanden und gleich L. Dann konvergiert* $\sum_{n=1}^{\infty} A_n e^{s_n x_0}$ *gegen den Wert L.*

Man kann ohne Beschränkung der Allgemeinheit $x_0=0$ annehmen. Aus der Zusatzbedingung folgt dann $c=-\frac{1}{p-q}$ und die absolute Konvergenz der Reihe für reelles $z=x>0$, da sich $|e^{s_n x}|$ wie $n^{-x(p-q)}$ verhält und A_n schneller als n^{-1} abnimmt. Der Beweis verläuft wie der des klassischen Tauberschen Satzes[4]). Es sei

$$S_n = \sum_{t=1}^{n} A_t,$$

so daß

$$|S_n - f(iy)| = \sum_{t=1}^{n} A_t(1-e^{s_t iy}) - \sum_{t=n+1}^{\infty} A_t e^{s_t iy} \leq \sum_{t=1}^{n} |A_t|\,|1-e^{s_t iy}| + \sum_{t=n+1}^{\infty} |A_t|\,|e^{s_t iy}|$$

wird. Da $Re(s_t iy)$ negativ ist, kann man wie oben im Anschluß an (12)

$$|1-e^{s_t iy}| \leq |s_t|\, y \leq C_{12}\, t\, y$$

schließen. Ferner ist

$$|e^{s_t iy}| = e^{-v_t y} \leq e^{-C_{13} t y}$$

mit passendem C_{13}. Es sei $M_n = \operatorname{Max}_{t>n} |A_t|$ und $M_n' = \operatorname{Max}_{t>n} |tA_t|$. Aus der Zusatzbedingung über die Koeffizienten folgt dann

$$M_n < M_n' n^{-1},$$

und man erhält

$$|S_n - f(iy)| \leq C_{12}\, y \sum_{t=1}^{n} t\,|A_t| + M_n' n^{-1} \sum_{t=n+1}^{\infty} |e^{-C_{13} t y}|.$$

Nun läßt man y die Folge n^{-1} durchlaufen. Dann ist wegen $t|A_t|\to 0$ auch

$$\lim_{n\to\infty} \sum_{t=1}^{n} \frac{t\,A|}{n} = 0.$$

[4]) LANDAU, E.: Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie. § 8.

Ferner ist

$$\lim_{n\to\infty} n^{-1}\sum_{t=n+1}^{\infty}|e^{-C_{13}ty}| = \lim_{\frac{1}{n}\to 0}\frac{\frac{1}{n}\cdot e^{-C_{13}\left(1+\frac{1}{n}\right)}}{1-e^{-C_{13}\cdot\frac{1}{n}}}$$

endlich, und da M_n' gegen Null geht, folgt $\lim_{n\to\infty}\left|S_n - f\left(\frac{i}{n}\right)\right| = \lim_{n\to\infty}|S_n - L| = 0$ und damit die Behauptung.

7. Existenz singulärer Punkte. Eine unmittelbare Folge aus Nr. **6** ist der Satz: *Es sei* $a=0$, $nA_n\to 0$ *und* $\sum_{n=1}^{\infty} A_n$ *divergent. Dann ist der Punkt* $z=0$ *singulär.* Andernfalls müßte ja $\lim f(iy)$ für $y\to 0$ existieren, und nach dem „Tauberschen" Satz wäre $\sum_{n=1}^{\infty} A_n$ konvergent. Aus der Funktionalgleichung (1) folgert man in diesem Fall auch die Singularität der Punkte $z=1, 2, \ldots$. Die Voraussetzung ist z. B. für $A_n = \frac{1}{n\log n}$ erfüllt.

8. Die Betrachtungen der letzten vier Nummern lassen sich auf Lösungen der Form

$$g(z) = \sum_{n=1}^{\infty} A_n' e^{s_n' z} \tag{16}$$

übertragen, worin s_n' die Nullstellen mit negativem Imaginärteil durchläuft. Diese Reihen konvergieren natürlich in „unteren" Halbebenen. Die Lösungen der Form (5) konvergieren, wenn sie überhaupt analytische Funktionen darstellen, in Horizontalstreifen; denn — vgl. [2]) — sie konvergieren für hinreichend große reelle t absolut, lassen sich daher als Summe einer Reihe (8) und einer Reihe (16) schreiben. Eine solche Summe ist also die allgemeinste Form einer analytischen Lösung der Funktionalgleichung (1). Ob die betrachteten Reihen über die Konvergenzhorizontalen analytisch fortsetzbar sind oder nicht — wobei auch die Möglichkeit besteht, daß die fortgesetzten Reihen nicht mehr der Funktionalgleichung genügen — läßt sich aus den abgeleiteten Sätzen nicht entnehmen.

Braunschweig, Mathematisches Institut der Technischen Hochschule.

(Eingegangen am 16. April 1955.)

GRUNSKY, H.
Math. Zeitschr. Bd. 63, S. 320—323 (1955)

Eine funktionentheoretische Integralformel.

Von
HELMUT GRUNSKY.

Meinem verehrten Lehrer ISSAI SCHUR zum Gedächtnis.

Der GAUSSsche Integralsatz kann als ein Analogon zu der Grundformel der Differential- und Integralrechnung angesehen werden: Wie diese das Integral über die Ableitung einer Funktion durch ihre Werte am Rande des Intervalles darstellt, so wird im GAUSSschen Integralsatz das Gebietsintegral über die Divergenz eines Vektorfeldes durch das Feld am Rande, d.h. jetzt durch ein Rand*integral*, ausgedrückt. Entsprechendes gilt für den STOKESschen Satz, der ja übrigens in dem uns hier allein interessierenden Fall der Ebene nichts Neues gegenüber dem GAUSSschen enthält. Es soll im folgenden gezeigt werden, daß sich in gewissen Fällen (auch wenn man von dem bekannten trivialen absieht, bei dem $\iint u_{xy}\,dx\,dy$ über ein achsenparalleles Rechteck betrachtet wird) die Integration noch einen Schritt weiter treiben läßt, so daß sich ein Gebietsintegral durch eine endliche Summe von Randwerten ausdrücken läßt. Das trifft nämlich zu für das Integral einer analytischen Funktion, erstreckt über ein Polygongebiet; es ist darstellbar durch eine Linearkombination der Werte der zweiten Stammfunktion in den Ecken mit Koeffizienten, die durch das Polygon allein bestimmt sind. Der einfache Beweis stützt sich auf eine Formulierung des GAUSS-STOKESschen Satzes mit den Darstellungsmitteln der komplexen Analysis, die zwar nicht neu[1]), aber doch nicht gerade geläufig ist. Ich beginne daher mit einem knappen Beweis dieser Formel.

Es sei $\mathfrak{f}(z)$ eine stetig differenzierbare, komplexwertige Funktion der komplexen Veränderlichen $z=x+iy$ in dem Gebiet $\mathfrak{G}$. Wir erklären auf dem System der BORELschen Teilmengen $\mathfrak{B}\subset\mathfrak{G}$ eine komplexwertige, totalstetige, abzählbar additive Mengenfunktion $\mathfrak{F}(\mathfrak{B})$, indem wir zunächst für ein beliebiges halboffenes Intervall $I\subset\mathfrak{G}: x_0\leq x<x_1,\ y_0\leq y<y_1$ setzen:

$$\mathfrak{F}(I)=\int\limits_{\mathrm{Rd}\,I}\mathfrak{f}(z)\,dz \tag{1}$$

und dann $\mathfrak{F}$ in bekannter Weise erweitern. Ist $\mathfrak{C}$ eine kompakte Teilmenge von $\mathfrak{G}$, deren Rand aus einem System C von endlich vielen einfach geschlossenen, rektifizierbaren Kurven besteht, so ist

$$\mathfrak{F}(\mathfrak{C})=\int\limits_{C}\mathfrak{f}(z)\,dz. \tag{2}$$

[1]) Vgl. z.B. BERGMANN, ST.: The kernel function and conformal mapping, p. 57 New York 1950.

$\mathfrak{F}(\mathfrak{B})$ hat eine Dichte, die wir ermitteln wollen. Wir schreiben die Formel für das totale Differential von $\mathfrak{f}(z)$ in der Form

$$(3) \qquad \mathfrak{f}(z) = \mathfrak{f}(z_0) + \mathfrak{f}_z(z_0)\cdot(z - z_0) + \mathfrak{f}_{\bar z}(z_0)\cdot(\bar z - \bar z_0) + o(|z - z_0|),$$

wobei die Koeffizienten $\mathfrak{f}_z(z_0)$ und $\mathfrak{f}_{\bar z}(z_0)$ eben durch diese Formel definiert sind. Die Umrechnung des in reellen Koordinaten geschriebenen totalen Differentials auf (3) ergibt den folgenden Zusammenhang mit $\mathfrak{f}_x = \frac{\partial \mathfrak{f}}{\partial x}$, $\mathfrak{f}_y = \frac{\partial \mathfrak{f}}{\partial y}$:

$$\mathfrak{f}_z = \frac{\partial \mathfrak{f}}{\partial z} = \frac{1}{2}\left(\frac{\partial \mathfrak{f}}{\partial x} - i\,\frac{\partial \mathfrak{f}}{\partial y}\right), \qquad \mathfrak{f}_{\bar z} = \frac{\partial \mathfrak{f}}{\partial \bar z} = \frac{1}{2}\left(\frac{\partial \mathfrak{f}}{\partial x} + i\,\frac{\partial \mathfrak{f}}{\partial y}\right).$$

Ist $\mathfrak{f}(z)$ speziell analytisch, so ersieht man aus (3) die identische Gültigkeit der Gleichung $\mathfrak{f}_{\bar z}(z) = 0$, und diese ist kennzeichnend für analytische Funktionen. Sie ist nichts anderes als eine besondere Schreibweise der CAUCHY-RIEMANNschen Differentialgleichungen. Ferner ist für ein analytisches $\mathfrak{f}(z)$:

$$\mathfrak{f}_z(z) = \mathfrak{f}'(z).$$

Setzen wir (3) in (1) ein, so ergeben die beiden ersten Glieder den Beitrag 0, das Restglied ein $o(|z_1 - z_0|^2)$, das dritte Glied dagegen

$$\mathfrak{f}_{\bar z}(z_0) \int_{\mathrm{Rd}\,I} \bar z\, dz = 2i\, I \cdot \mathfrak{f}_{\bar z}(z_0),$$

wo I jetzt den Inhalt des Intervalles bedeutet. Läßt man nun z_1 nach z_0 rücken, so erkennt man, daß die Dichte von $\mathfrak{F}(\mathfrak{B})$ gleich $2i\,\mathfrak{f}_{\bar z}$ ist. Es folgt unmittelbar:

$$(4) \qquad \int_C \mathfrak{f}(z)\, dz = 2i \iint_{\mathfrak{C}} \frac{\partial \mathfrak{f}}{\partial \bar z}\,(dz),$$

wo (dz) ein Flächenelement bezeichnet und C und $\mathfrak{C}$ dieselbe Bedeutung wie in (2) haben. — Diese Formel geht für analytische $\mathfrak{f}(z)$ in den CAUCHYschen Integralsatz über. Zerlegt man (4) in Real- und Imaginärteil, so findet man den STOKESschen und den GAUSSschen Integralsatz für das Vektorfeld $\overline{\mathfrak{w}} = \overline{\mathfrak{f}(z)}$.

Gehen wir in (4) beiderseits zum konjugiert Komplexen über und schreiben statt $\overline{\mathfrak{f}(z)}$ wieder $\mathfrak{f}(z)$, so erhalten wir:

$$\iint_{\mathfrak{C}} \frac{\partial \mathfrak{f}}{\partial z}\,(dz) = \frac{i}{2}\int_C \mathfrak{f}(z)\, d\bar z.$$

Wir wollen nun diese Formel insbesondere auf eine analytische Funktion $\mathfrak{f}(z)$ anwenden. Setzen wir noch $\mathfrak{f}(z) = f'(z)$ so bekommen wir:

$$(5) \qquad \iint_{\mathfrak{C}} f''(z)\,(dz) = \frac{i}{2}\int_C f'(z)\, d\bar z.$$

Das Integral der rechten Seite läßt sich bei besonderen Voraussetzungen über C auswerten. Es sei insbesondere $\mathfrak{C} = \mathfrak{P}$ ein ein- oder mehrfach zusammenhängendes Polygongebiet mit Rand P und den Ecken z_ν, $\nu = 1, \ldots, n$; $z_{\nu+}$ sei

die bei einem positiven Umlauf auf z_ν folgende, $z_{\nu-}$ die vorangehende Ecke, P_ν die Seite $(z_\nu, z_{\nu+})$, deren Richtung durch ε_ν mit $|\varepsilon_\nu| = 1$ gekennzeichnet werde: $z_{\nu+} - z_\nu = \varepsilon_\nu |z_{\nu+} - z_\nu|$. Dann ist:

$$\int_{P_\nu} f'(z)\, d\bar{z} = \varepsilon_\nu^{-2} \int_{P_\nu} f'(z)\, dz = \varepsilon_\nu^{-2} \big(f(z_{\nu+}) - f(z_\nu)\big).$$

Summiert man über alle Polygonseiten und ordnet nach den $f(z_\nu)$, so drückt sich die rechte Seite von (5) als Linearverbindung dieser Größen aus, und man erhält:

$$\iint_{\mathfrak{P}} f''(z)\, (dz) = \sum_{\nu=1}^{n} \delta_\nu f(z_\nu). \tag{6}$$

mit

$$\delta_\nu = \frac{i}{2} \left(\varepsilon_{\nu-}^{-2} - \varepsilon_\nu^{-2}\right). \tag{7}$$

Wesentlich ist, daß die δ_ν durch das Polygon allein bestimmt und von $f(z)$ unabhängig sind.

Ein einfacher Grenzübergang erlaubt es schließlich, die Voraussetzung, daß $f(z)$ auch auf dem Rande des Polygongebietes regulär sei, durch die der Stetigkeit im abgeschlossenen Bereich zu ersetzen. — Somit ist bewiesen:

Satz. *$f(z)$ sei regulär in dem Gebiet des geradlinig berandeten, ein- oder mehrfach zusammenhängenden Polygons $\mathfrak{P}$ mit den Ecken z_ν $(\nu = 1, \ldots, n)$ und stetig in dem zugehörigen abgeschlossenen Bereich. Dann existiert die linke Seite von* (6) *und es gilt* (6) *mit* (7), *wo ε_ν mit $|\varepsilon_\nu| = 1$ die Richtung der von z_ν ausgehenden, $\varepsilon_{\nu-}$ die der in z_ν mündenden Polygonseite kennzeichnet.*

Die Analogie dieses Ergebnisses mit der Grundformel der Differential- und Integralrechnung ist offenkundig. Diese Analogie legt viele Anwendungsmöglichkeiten nahe, auf die ich an anderer Stelle eingehe[2]).

Von besonderer Wichtigkeit ist natürlich der Fall, daß $\mathfrak{P} = \mathfrak{D}$ ein Dreieck ist. Für die δ_ν gibt es dann eine andere zweckmäßige Darstellung, die man leicht unabhängig von (7) findet, wenn nur überhaupt die Gültigkeit einer Formel (6) bekannt ist. Wenden wir nämlich diese nacheinander auf $f(z) = 1$, $f(z) = z$, $f(z) = z^2$ an, so erhalten wir für die δ_ν, $\nu = 1, 2, 3$ die drei Gleichungen:

$$\sum \delta_\nu = 0, \quad \sum z_\nu \delta_\nu = 0, \quad \sum z_\nu^2 \delta_\nu = 2\mathfrak{D}$$

und damit

$$\delta_\nu = \frac{2\mathfrak{D}}{(z_\nu - z_{\nu-1})(z_\nu - z_{\nu+1})}, \tag{8}$$

wo die Indizes mod 3 zählen. Natürlich bestätigt man leicht die Übereinstimmung von (7) und (8).

Setzt man nun noch voraus $f(z_1) = f(z_2) = 0$ und schreibt z für z_3, für die Integrationsveränderliche dagegen ζ, so folgt als Darstellung der in zwei

[2]) Über eine Anwendung auf dem Gebiete der Differenzenrechnung habe ich auf dem Internationalen Mathematikerkongreß 1954, Amsterdam, berichtet. Vgl. Proc. Int. Congr. Math. 2, 114—115 (1954).

gegebenen Punkten verschwindenden zweiten Stammfunktion zu einer gegebenen analytischen Funktion $f''(z)$:

$$f(z) = \frac{1}{2}(z - z_1)(z - z_2)\,\mathfrak{M}(f''(\zeta); z), \tag{9}$$

wo $\mathfrak{M}$ den Mittelwert von f'' über das Dreieck $\mathfrak{D}$ mit den Ecken z_1, z_2, z bedeutet:

$$\mathfrak{M}(f''(\zeta); z) = \frac{1}{\mathfrak{D}} \iint_{\mathfrak{D}} f''(\zeta)\,(d\zeta).$$

Falls z auf der Verbindungsstrecke von z_1 und z_2 liegt und diese Punkte (ohne Beschränkung der Allgemeinheit) als reell angenommen werden, so geht (9) in die bekannte Integraldarstellung einer Funktion $f(z)$ mit $f(z_1) = f(z_2) = 0$ mittels $f''(z)$ über:

$$f(z) = \int_{z_1}^{z_2} k(z, \zeta)\, f''(\zeta)\, d\zeta$$

mit

$$k(z, \zeta) = \begin{cases} \dfrac{(\zeta - z_1)(z - z_2)}{z_2 - z_1} & \zeta \leq z \\ \dfrac{(\zeta - z_2)(z - z_1)}{z_2 - z_1} & \zeta \geq z. \end{cases}$$

Aus (9) ergibt sich als unmittelbare Folgerung: Ist $f(z)$ regulär in einem konvexen Gebiet $\mathfrak{G}$ und ist $f(z_1) = f(z_2) = 0$, $|f''(z)| \leq 1$, so gilt in $\mathfrak{G}$:

$$|f(z)| \leq \frac{1}{2}|z - z_1| \cdot |z - z_2|.$$

Mainz, Mathematisches Institut der Universität.

(Eingegangen am 28. April 1955.)

SPECHT, W.
Math. Zeitschr. Bd. 63, S. 324—330 (1955)

Abschätzung der Wurzeln algebraischer Gleichungen. III.

ISSAI SCHUR zum Gedächtnis.

Von

WILHELM SPECHT.

Bestimmt man zu einem komplexen Polynom

$$f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n \tag{1}$$

des Grades n die Größen

$$\alpha = \max_{1 \le \nu \le n} \sqrt[\nu]{a_\nu}\,, \tag{2.1}$$

$$\beta = \max_{1 \le \nu \le n} |a_\nu|\,, \tag{2.2}$$

so bestehen für die nach ihrem absoluten Betrage

$$|\xi_1| \ge |\xi_2| \ge \cdots \ge |\xi_n| \tag{3}$$

geordneten Nullstellen $\xi_1, \xi_2, \ldots, \xi_n$ des Polynoms Ungleichungen[1])

$$|\xi_1 \xi_2 \ldots \xi_k| < \sigma_k \alpha^k \quad \text{für} \quad 1 \le k \le n \tag{4}$$

mit den nur von der Stellenzahl k, nicht aber vom Grade n abhängigen Konstanten

$$\sigma_k = k^{-\frac{k}{2}} (k+1)^{\frac{k+1}{2}} \quad \text{für} \quad 1 \le k < \infty \tag{4*}$$

und[2])

$$|\xi_1 \xi_2 \ldots \xi_k| < 1 + k\beta \quad \text{für} \quad 1 \le k \le n. \tag{5}$$

Für die ersten Ungleichungen (4) *soll hier ein einfacherer Beweis als früher gegeben werden, der zugleich unter Berücksichtigung des Grades n zur genaueren Aussage*

$$|\xi_1 \xi_2 \ldots \xi_k| \le \sigma_k \alpha^k \left[1 - \left(\frac{k}{k+1}\right)^{n+1}\right]^{\frac{1}{2}} \quad \text{für} \quad 1 \le k \le n \tag{4**}$$

führt. Die zweiten Ungleichungen (5) *lassen sich zu den Abschätzungen*

$$|\xi_1 \xi_2 \ldots \xi_k| < \tau_k(\beta) \quad \text{für} \quad 1 \le k \le n \tag{5*}$$

[1]) Vgl. SPECHT, W.: Abschätzung der Wurzeln algebraischer Gleichungen. Math. Z. 52, 310—321 (1949), Satz 4.

[2]) Vgl. SPECHT, W.: Zur Theorie der algebraischen Gleichungen. Jahresber. D.M.V. 48, 142—145 (1938).

mit den gleichfalls vom Grade n unabhängigen Konstanten

$$(5^{**}) \qquad \begin{cases} \tau_k(\beta) = \left(\dfrac{(k-1)\beta^2 + \beta\sqrt{(k-1)^2\beta^2 + 4k} + 2}{2}\right)^{\frac{1}{2}} \times \\ \qquad \times \left(\dfrac{2k - (k-1)\beta^2 + \beta\sqrt{(k-1)^2\beta^2 + 4k}}{2k}\right)^{k/2} \end{cases}$$

verbessern. Ohne Einschränkung darf für das Weitere

$$\alpha > 0 \quad \text{und} \quad \beta > 0$$

vorausgesetzt werden.

Als gemeinsame Grundlage dient folgende Bemerkung[3]):

Mit der Norm

$$N = N(f) = (1 + |a_1|^2 + |a_2|^2 + \cdots + |a_n|)^{\frac{1}{2}}$$

des Polynoms $f(x)$ *gilt*

$$|\xi_1 \xi_2 \ldots \xi_k| \leq N \quad \text{für} \quad 1 \leq k \leq n.$$

Setzen wir demgemäß mit einer vorerst unbestimmten positiven Größe u

$$y = u\,x \quad \text{und} \quad g(y) = u^n f(x) = y^n + a_1 u\, y^{n-1} + a_2 u^2 y^{n-2} + \cdots + a_n u^n,$$

so gelten für die Nullstellen

$$\eta_\nu = u\,\xi_\nu \quad \text{für} \quad 1 \leq \nu \leq n$$

des Polynoms $g(y)$ die Ungleichungen

$$|\eta_1 \eta_2 \cdots \eta_k|^2 \leq 1 + |a_1 u|^2 + |a_2 u^2|^2 + \cdots + |a_n u^n|^2,$$

sodaß für die Nullstellen des Polynoms $f(x)$ mit einem beliebigen Wert $z = u^2$ folgt:

$$(6) \qquad |\xi_1 \xi_2 \ldots \xi_k|^2 \leq \frac{1 + |a_1|^2 z + |a_2|^2 z^2 + \cdots + |a_n|^2 z^n}{z^k} \quad \text{für} \quad 1 \leq k \leq n.$$

Unter der Voraussetzung

$$|a_\nu| \leq \alpha^\nu \quad \text{für} \quad 1 \leq \nu \leq n \quad \text{mit} \quad \alpha = \max_{1 \leq \nu \leq n} \sqrt[\nu]{a_\nu}$$

erhalten wir aus (6)

$$|\xi_1 \xi_2 \ldots \xi_k|^2 \leq \frac{1 + \alpha^2 z + \alpha^4 z^2 + \cdots + \alpha^{2n} z^n}{z^k}$$

und unter der Beschränkung

$$0 < \alpha^2 z = t < 1$$

die Ungleichung

$$(7) \qquad |\xi_1 \xi_2 \ldots \xi_k|^2 \leq \alpha^{2k} \frac{1 - t^{n+1}}{(1-t)\,t^k}$$

[3]) Vgl. a. a. O. [1]), Satz 1.

und ohne Berücksichtigung des Grades n des Polynoms $f(x)$ die Ungleichung

$$(7^*) \qquad |\xi_1 \xi_2 \dots \xi_k|^2 < \frac{\alpha^{2k}}{(1-t)\,t^k}.$$

Die Funktion

$$F(t) = (1-t)\,t^k$$

nimmt ihr Maximum an der Nullstelle $t_0 = \frac{k}{k+1}$ ihrer Ableitung

$$F'(t) = k\,t^{k-1} - (k+1)\,t^k$$

an; für die Stelle t_0 gilt

$$F(t_0) = (1-t_0)\,t_0^k = \left(1 - \frac{k}{k+1}\right)\left(\frac{k}{k+1}\right)^k = \frac{k^k}{(k+1)^{k+1}}.$$

Damit gewinnen wir in der Tat

$$|\xi_1 \xi_2 \dots \xi_k|^2 < k^{-k}(k+1)^{k+1}\alpha^{2k} \quad \text{für} \quad 1 \leq k \leq n,$$

wie in (4) behauptet worden war.

Bei Berücksichtigung des Grades n des Polynoms $f(x)$ entnimmt man der Ungleichung (7) für den gleichen Wert $t_0 = \frac{k}{k+1}$ des Parameters t die Abschätzung

$$|\xi_1 \xi_2 \dots \xi_k|^2 \leq k^{-k}(k+1)^{k+1}\left(1 - \left(\frac{k}{k+1}\right)^{n+1}\right)\alpha^{2k},$$

in Übereinstimmung mit der Ungleichung (4**).

Unter Verwendung des Parameterwertes

$$\alpha^2 z_0 = t_0 \quad \text{oder} \quad z_0 = \alpha^{-2}\,\frac{k}{k+1}$$

können wir bei der genaueren Festsetzung

$$a_\nu = \alpha^\nu \varepsilon_\nu \quad \text{mit} \quad |\varepsilon_\nu| \leq 1 \quad \text{und} \quad \alpha = \max \sqrt[\nu]{a_\nu} \quad \text{für} \quad 1 \leq \nu \leq n$$

die Ungleichung (6) auswerten zu der Abschätzung

$$|\xi_1 \xi_2 \dots \xi_k|^2 \leq$$

$$\leq \alpha^{2k}\left(\frac{k}{k+1}\right)^{-k}\left(1 + |\varepsilon_1|^2\,\frac{k}{k+1} + |\varepsilon_2|^2\left(\frac{k}{k+1}\right)^2 + \dots + |\varepsilon_n|^2\left(\frac{k}{k+1}\right)^n\right),$$

also

$$(8) \qquad |\xi_1 \xi_2 \dots \xi_k| \leq \alpha^k \left[\frac{(k+1)^n + |\varepsilon_1|^2(k+1)^{n-1}k + |\varepsilon_2|^2(k+1)^{n-2}k^2 + \dots + |\varepsilon_n|^2 k^n}{(k+1)^{n-k}k^k}\right]^{\frac{1}{2}}.$$

Für die Nullstelle ξ_1 maximalen absoluten Betrages folgt hieraus insbesondere

$$(8^*) \qquad |\xi_1| \leq 2\alpha\left[\frac{2^n + |\varepsilon_1|^2 2^{n-1} + |\varepsilon_2|^2 2^{n-2} + \dots + |\varepsilon_n|^2}{2^{n+1}}\right] < 2\alpha.$$

Die aus der Ungleichung

$$(6) \qquad |\xi_1 \xi_2 \dots \xi_k|^2 \leq \frac{1 + |a_1|^2 z + |a_2|^2 z^2 + \dots + |a_n|^2 z^n}{z^k} = \frac{g(z)}{z^k} = \Phi_k(z)$$

hervorgehende günstigste Schranke

$$|\xi_1 \xi_2 \dots \xi_k| \leq \Phi_k^{\frac{1}{2}}(z^*)$$

liefert natürlich die (einzige) positive Nullstelle z^* der Ableitung

$$\Phi_k'(z) = \frac{z\,g'(z) - k\,g(z)}{z^{k+1}},$$

also der Gleichung

$$\varphi(z) = k + \sum_{\nu=1} (k - \nu)\,|a_\nu|^2 z^\nu = 0.$$

Auf die Ungleichung (4) stützt sich auch die folgende Bemerkung über reelle Polynome

$$f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n \quad \text{mit} \quad \alpha = \max_{1 \leq \nu \leq n} |\sqrt[\nu]{a_\nu}|.$$

Für jedes Paar $\omega, \overline{\omega}$ konjugierter imaginärer Nullstellen des reellen Polynoms $f(x)$ besteht die Ungleichung

$$|\omega\overline{\omega}| < \sigma_2 \alpha^2 \quad \text{oder} \quad |\omega| < \sigma_2^{\frac{1}{2}} \alpha.$$

Eine Nullstelle ξ_1 des reellen Polynoms $f(x)$, die dem Kreisring

$$\sigma_2^{\frac{1}{2}} \alpha \leq |z| < \sigma_1 \alpha$$

der Zahlenebene angehört, ist demnach reell und durch diese Begrenzung eindeutig bestimmt.

Liegt dieser Sonderfall vor, so erfüllen alle Paare $\omega, \overline{\omega}$ konjugierter imaginärer Nullstellen die Ungleichung

$$\sigma_2^{\frac{1}{2}} \alpha\,|\omega\overline{\omega}| \leq |\xi_1 \xi_2 \xi_3| < \sigma_3 \alpha^3 \quad \text{oder} \quad |\omega| < \sigma_3^{\frac{1}{2}} \sigma_2^{-\frac{1}{2}} \alpha.$$

Setzt man demgemäß

$$\tau_0 = \sigma_1 = 2, \quad \tau_1 = \sigma_2^{\frac{1}{2}} = 2^{-\frac{1}{2}} 3^{\frac{3}{2}} \sim 1{,}612, \quad \tau_2 = \sigma_3^{\frac{1}{2}} \sigma_2^{-\frac{1}{2}} = 2^{\frac{3}{2}} 3^{-\frac{3}{2}} \sim 1{,}382,$$

so gilt:

Ist $f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n$ ein reelles Polynom mit der Schranke

$$\alpha = \max_{1 \leq \nu \leq n} |\sqrt[\nu]{a_\nu}|,$$

so liegen alle Nullstellen ξ des Polynoms im Kreisbereich

$$0 \leq |z| < \tau_0 \alpha$$

und mit höchstens einer (reellen) Ausnahme ξ_1 im Kreisbereich

$$0 \leq |z| < \tau_1 \alpha.$$

Liegt eine Nullstelle ξ_1 im Kreisbereich

$$\tau_1 \alpha \leq |z| < \tau_0 \alpha,$$

so liegen alle weiteren reellen bzw. imaginären Nullstellen im Kreisbereich

$$0 \leq |z| < \tau_1 \alpha \quad \text{bzw.} \quad 0 \leq |z| < \tau_2 \alpha.$$

Bei Verwendung der in (2.2) erklärten Größe β erhalten wir aus (6) die Ungleichung

$$|\xi_1 \xi_2 \dots \xi_k|^2 \leq \frac{1 - z + \beta^2 (z - z^{n+1})}{(1 - z) z^k} \tag{9}$$

und ohne Berücksichtigung des Grades n des Polynoms $f(x)$ auch

$$|\xi_1 \xi_2 \dots \xi_k|^2 < \frac{1 + (\beta^2 - 1) z}{(1 - z) z^k} = G(z). \tag{9*}$$

Um das Minimum der Funktion $G(z)$ im Intervall $0 < z < 1$ zu bestimmen, bilden wir mit der Abkürzung $\delta = \beta^2 - 1$ die Ableitung

$$G'(z) = \frac{(1 - z)\,\delta z - (k - (k + 1) z)(1 + \delta z)}{(1 - z)^2 z^{k+1}} = \frac{g(z)}{(1 - z)^2 z^{k+1}}$$

mit dem Polynom

$$g(z) = k \delta z^2 - \big((k - 1)\,\delta - (k + 1)\big) z - k,$$

das wegen

$$g(0) = -k; \quad g(1) = \delta + 1 = \beta^2 > 0$$

genau eine Nullstelle z_0 im Intervall $0 < z_0 < 1$ besitzt, nämlich

$$z_0 = \begin{cases} \dfrac{k}{k + 1} & \text{für} \quad \beta = 1, \\ \dfrac{(k - 1)\beta^2 - 2k + \beta \sqrt{(k - 1)^2 \beta^2 + 4k}}{2(\beta^2 - 1) k} & \text{für} \quad \beta \neq 1. \end{cases}$$

Daraus folgt

$$G(z_0) = \frac{(k + 1)^{k+1}}{k^k}$$

bzw.

$$G(z_0) = (\beta^2 - 1) \frac{(k - 1)\beta + \sqrt{(k - 1)^2 \beta^2 + 4k}}{(k + 1)\beta - \sqrt{(k - 1)^2 \beta^2 + 4k}} \left(\frac{2(\beta^2 - 1) k}{(k - 1)\beta^2 - 2k + \beta \sqrt{(k - 1)^2 \beta^2 + 4k}} \right)^k.$$

Setzt man im Falle $\beta \neq 1$ abkürzend

$$R = \sqrt{(k - 1)^2 \beta^2 + 4k},$$

so gewinnt man bei Berücksichtigung der Gleichungen

$$[(k + 1)\beta - R]\,[(k + 1)\beta + R] = 4k(\beta^2 - 1),$$

$$[(k - 1)\beta + R]\,[(k + 1)\beta + R] = 2k[(k - 1)\beta^2 + 2 + \beta R],$$

$$[(k - 1)\beta^2 - 2k + \beta R]\,[(k - 1)\beta^2 - 2k - \beta R] = -4k^2(\beta^2 - 1)$$

die einfachere Gestalt

$$G(z_0) = \frac{(k-1)\beta^2 + \beta R + 2}{2}\left(\frac{2k - (k-1)\beta^2 + \beta R}{2k}\right)^k$$

und damit die Ungleichung

$$|\xi_1 \xi_2 \dots \xi_k| < G^{\frac{1}{2}}(z_0) = \tau_k(\beta) \quad \text{für} \quad 1 \leq k \leq n \tag{10}$$

mit den in (5**) angegebenen Werten; sie ist übrigens auch im Falle $\beta = 1$ gültig.

Daß die Ungleichungen (10) die Abschätzungen (5) verbessern, geht aus den Beziehungen

$$\tau_1(\beta) = 1 + \beta \quad \text{für} \quad 0 < \beta; \tag{11.1}$$

$$\tau_k(\beta) < 1 + k\beta \quad \text{für} \quad 0 < \beta;\ 2 \leq k < \infty \tag{11.2}$$

hervor. Zunächst stellen wir die Gleichung

$$\tau_1(\beta) = 1 + \beta$$

leicht fest, so daß weiterhin $k \geq 2$ angenommen werden darf. Auf Grund der Minimaleigenschaft des Wertes z_0 bestehen für die Parameterwerte

$$0 < z_1 = \frac{k}{k+\beta} < 1 \quad \text{und} \quad 0 < z_2 = \frac{k}{k+1} < 1$$

die Beziehungen

$$\tau_k^2(\beta) = G(z_0) \leq G(z_1) = (1 + \beta k)\left(\frac{k+\beta}{k}\right)^k < (1 + \beta k)\, e^\beta \tag{12}$$

bzw.

$$\tau_k^2(\beta) = G(z_0) \leq G(z_2) = (1 + \beta^2 k)\left(\frac{k+1}{k}\right)^k < (1 + \beta^2 k)\, e. \tag{12*}$$

Im Falle $0 < \beta < 1$ gilt

$$0 < \frac{e^\beta - 1}{\beta} < e - 1 < 2 \leq k, \quad \text{also} \quad e^\beta < 1 + \beta k$$

und damit nach (12)

$$\tau_k^2(\beta) < (1 + k\beta)\, e^\beta < (1 + k\beta)^2, \quad \text{also} \quad \tau_k(\beta) < 1 + k\beta.$$

Im Falle $1 \leq \beta$ und $3 \leq k$ entnehmen wir aus (12*)

$$\tau_k^2(\beta) < (1 + \beta^2 k)\, e < k + \beta^2 k^2 < (1 + \beta k)^2.$$

Im allein noch verbleibenden Falle $k = 2$ und $\beta \geq 1$ entnehmen wir aus (12*) zunächst

$$\tau_2^2(\beta) < (1 + 2\beta^2)\left(\tfrac{3}{2}\right)^2 = \tfrac{9}{4} + \tfrac{9}{2}\beta^2.$$

Die Ungleichung

$$\tfrac{9}{4} + \tfrac{9}{2}\beta^2 \geq (1 + 2\beta)^2 = 1 + 4\beta + 4\beta^2$$

besteht wegen $\beta \geq 1$ nur im Falle

$$\beta \geq 4 + \tfrac{3}{2}\sqrt{6} > 7.$$

Andererseits erhalten wir aus

$$\tau_2^2(\beta) = \left(\frac{\beta^2 + \beta\sqrt{\beta^2+8} + 2}{2}\right)\left(\frac{4 - \beta^2 + \beta\sqrt{\beta^2+8}}{4}\right)^2$$

wegen

$$\sqrt{\beta^2+8} < \beta + \frac{4}{\beta}$$

die Abschätzung

$$\tau_2^2(\beta) < (\beta^2+3)\,4.$$

Die Ungleichung

$$(\beta^2+3)\,4 \geq (1+2\beta)^2 = 1 + 4\beta + 4\beta^2$$

besteht nur im Falle

$$12 \geq 1 + 4\beta \quad \text{oder} \quad \beta \leq \frac{11}{4}.$$

Mithin gilt auch

$$\tau_2(\beta) < 1 + 2\beta \quad \text{für} \quad \beta \geq 1.$$

Erlangen, Mathematisches Institut der Universität.

(Eingegangen am 3. Juli 1955.)

PÓLYA, G.
Math. Zeitschr. Bd. 63, S. 331—337 (1955)

On the characteristic frequencies of a symmetric membrane *).

Dedicated to the memory of I. SCHUR.

By

G. PÓLYA.

1. *Notation.* We consider a simply connected open domain $\mathfrak{D}$ in the z-plane, $z=x+iy$; we let $\mathfrak{C}$ denote the (sufficiently smooth) boundary curve of $\mathfrak{D}$. The function $u=u(x, y)$ is defined in $\mathfrak{D}$ and satisfies the differential equation

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\nu^2 u=0 \quad \text{in } \mathfrak{D}. \tag{1.1}$$

We consider two, very different, boundary conditions:

$$u=0 \quad \text{on } \mathfrak{C}, \tag{1.2}$$

$$\frac{\partial u}{\partial n}=0 \quad \text{on } \mathfrak{C}; \tag{1.3}$$

n stands for the normal to $\mathfrak{C}$. In both cases, there exists an infinity of characteristic frequencies ν; let $\lambda_1, \lambda_2, \ldots$ denote the values of ν with the condition (1.2), and $\mu_1, \mu_2, \ldots$ with the condition (1.3);

$$0<\lambda_1<\lambda_2\leq\lambda_3\leq\cdots, \tag{1.4}$$

$$0=\mu_1<\mu_2\leq\mu_3\leq\cdots. \tag{1.5}$$

Let $\dot{r}$ denote the *maximum inner conformal radius of* $\mathfrak{D}$, defined as follows. We map the interior of $\mathfrak{D}$ conformally onto the interior of a circle so that the point a of $\mathfrak{D}$ corresponds to the center of the circle and the linear magnification at the point a is equal to 1; let the radius of the circle so obtained be r_a. The value of r_a varies with the point a, and $\dot{r}$ is the maximum of r_a.

We say that the domain $\mathfrak{D}$ possesses *symmetry of order* n if $\mathfrak{D}$ coincides with itself after a rotation through an angle $2\pi/n$ about a certain point O, called the center of symmetry if $n\geq 2$. Any domain possesses symmetry of order 1, a square has symmetry of order 4, a circle symmetry of every order.

2. *Aim.* The products $\lambda_n \dot{r}$ and $\mu_n \dot{r}$ depend only on the shape of the domain $\mathfrak{D}$, not on its size. G. SZEGÖ and the present author have shown that $\lambda_1 \dot{r}$ attains its maximum when $\mathfrak{D}$ becomes a circle [3, p. 97—98][1]). M. SCHIFFER and the present author have shown that, for any n,

$$\left(\frac{1}{\lambda_1^2}+\frac{1}{\lambda_2^2}+\cdots+\frac{1}{\lambda_n^2}\right)\frac{1}{\dot{r}^2}$$

*) Sponsored by the Office of Naval Research.

[1]) Numerals in square brackets refer to the Bibliography at the end of the paper.

attains its minimum when $\mathfrak{D}$ becomes a circle [2, p. 304—307; attention to the difference of notation]. It follows from a theorem recently proved by G. SZEGÖ that also $\mu_2 \dot{r}$ attains its maximum for the circle [4]. Yet, for higher values of n, we can not generally assert that $\lambda_n \dot{r}$ or $\mu_n \dot{r}$ attains its maximum for the circle. In fact, Table 1 shows that both assertions are wrong for $n = 5$.

If, however, we restrict the choice of $\mathfrak{D}$ and admit only convex domains with sufficiently high symmetry, both $\lambda_n \dot{r}$ and $\mu_n \dot{r}$ will be maximum for the circle. This is expressed more precisely by the following theorem.

Table 1.

n	$\lambda_n \dot{r}$		$\mu_n \dot{r}$	
	Circle	Square	Circle	Square
1	2.405	2.396	0.000	0.000
2	3.832	3.789	1.841	1.695
3	3.832	3.789	1.841	1.695
4	5.135	4.793	3.054	2.396
5	5.135	5.358	3.054	3.389

Table 2.

n	$2n+1$	$l(n)$	$m(n)$
1	3	3	3
2	5	6	6
3	7	10	8
4	9	15	12
5	11	19	17

There are two increasing functions $l(n)$ and $m(n)$ of the positive integer n, with the following property: If the domain $\mathfrak{D}$ is convex and possesses a symmetry of order not lower than $2n+1$ but is not a circle, each of the $l(n)+m(n)-1$ quantities

$$\lambda_1 \dot{r}, \lambda_2 \dot{r}, \ldots, \lambda_{l(n)} \dot{r}, \quad \mu_2 \dot{r}, \ldots, \mu_{m(n)} \dot{r}$$

belonging to $\mathfrak{D}$ will be inferior to the corresponding quantity belonging to a circle.

Table 2 contains a few values of $l(n)$ and $m(n)$.

As it can be seen from the full definition in Section 4 (a)

$$l(n+1) - l(n) \geq 2, \quad m(n+1) - m(n) \geq 2. \tag{2.1}$$

For example, let us consider a square. It has symmetry of order 4, which is higher than 3. Now, as Table 2 shows, $l(1) = m(1) = 3$, and so our theorem asserts that each of the values

$$\lambda_1 \dot{r}, \lambda_2 \dot{r}, \lambda_3 \dot{r}, \mu_2 \dot{r}, \mu_3 \dot{r}$$

belonging to the square must be less than the corresponding value belonging to the circle; see Table 1. Yet for $\lambda_5 \dot{r}$ or $\mu_5 \dot{r}$ such an assertion is wrong, as Table 1 shows.

The principal aim of the present paper is to prove the theorem here formulated. I stated that part of the theorem that is concerned with λ_n some time ago [1]. Additional remarks are presented in Section 5.

3. *Lemma.* It is convenient to define the frequencies $\lambda_1, \lambda_2, \ldots, \mu_1, \mu_2, \ldots$ by their extremal properties which are connected with the RAYLEIGH ratio

$$R[f] = \frac{\iint |\operatorname{grad} f|^2 \, dx \, dy}{\iint f^2 \, dx \, dy}. \tag{3.1}$$

The function $f=f(z)$ can be considered also as a function of the two real variables x and y; f is defined in the domain $\mathfrak{D}$, which is also the domain of integration of both double integrals arising in (3.1); of course

$$|\operatorname{grad} f|^2 = f_x^2 + f_y^2.$$

The most convenient is, in our case as in many other cases, to use a lemma due to POINCARÉ [2, p. 275–278] which we formulate here in detail only for one of the two cases with which we are concerned.

Suppose that the functions $g_1, g_2, \ldots, g_l$, defined, realvalued, and sufficiently smooth in $\mathfrak{D}$ and on $\mathfrak{C}$, are linearly independent and that

$$g_1 = g_2 = \cdots = g_l = 0 \text{ on } \mathfrak{C}. \tag{3.2}$$

Suppose also that the quadratic forms $\sum\sum a_{ik}x_ix_k$ and $\sum\sum b_{ik}x_ix_k$, the latter of which is positive definite, are such that for all values of $x_1, x_2, \ldots, x_l$ (0, 0, …, 0, of course, excepted)

$$R[x_1g_1 + x_2g_2 + \cdots + x_lg_l] \leq \frac{\sum\sum a_{ik}x_ix_k}{\sum\sum b_{ik}x_ix_k}. \tag{3.3}$$

Let the roots of the characteristic equation

$$\begin{vmatrix} a_{11} - x\,b_{11} \ldots a_{1l} - x\,b_{1l} \\ \cdots\cdots\cdots\cdots \\ a_{l1} - x\,b_{l1} \ldots a_{ll} - x\,b_{ll} \end{vmatrix} = 0 \tag{3.4}$$

be $x = \lambda_1'^2, \lambda_2'^2, \ldots, \lambda_l'^2$, increasingly ordered so that

$$0 < \lambda_1' \leq \lambda_2' \leq \cdots \leq \lambda_l'. \tag{3.5}$$

Then

$$\lambda_1 \leq \lambda_1', \quad \lambda_2 \leq \lambda_2', \ldots, \lambda_l \leq \lambda_l'. \tag{3.6}$$

It is convenient to call $g_1, \ldots, g_l$ the "trial functions". We shall need two corollaries of this lemma.

(a) *If the case of equality is excluded in* (3.3), *it is necessarily also excluded in each of the l resulting relations* (3.6). The proof given l. c. [2, p. 277] yields also this corollary.

(b) *If*

$$a_{ik} = \iint \operatorname{grad} g_i \cdot \operatorname{grad} g_k \, dx\, dy$$
$$b_{ik} = \iint g_i g_k \, dx\, dy$$

(so that the inequality (3.3) *becomes an equation) and the trial functions $g_1, g_2, \ldots, g_l$ coincide with the first l eigenfunctions of $\mathfrak{D}$, all the l inequalities* (3.6) *become equations.* Also this corollary is proved l. c. [2, p. 278].

4. *Proof.* It will be enough to prove what the theorem of Section 2 states about λ_n; the proof of the similar assertion dealing with μ_n is strictly parallel and presents no additional difficulty.

(a) The eigenfunctions of the equation (1.1) under the boundary condition (1.2) are well known when the domain is the unit circle. They are of the form

$$J_p(\lambda\varrho)\cos p\vartheta, \quad J_p(\lambda\varrho)\sin p\vartheta \tag{4.1}$$

where ϱ and ϑ are the polar coordinates of an arbitrary point ζ inside the unit circle; that is,

$$\zeta = \varrho e^{i\vartheta} \tag{4.2}$$

with $0 \leq \varrho < 1$, $-\pi \leq \vartheta \leq \pi$. The corresponding characteristic frequency is λ, and p is an integer that we shall call the *type* of the eigenfunction. If the type is 0, only the first of the two expressions (4.1) is an eigenfunction, the second vanishes identically.

We define $l(n)$ as the greatest integer with the property that none of the eigenfunctions belonging to the characteristic frequencies

$$\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{l(n)} \tag{4.3}$$

is of a type superior to n. We call the eigenfunctions corresponding to the frequencies (4.3)

$$\chi_1(\zeta), \chi_2(\zeta), \chi_3(\zeta), \ldots, \chi_{l(n)}(\zeta), \tag{4.4}$$

respectively.

(b) We consider the domain $\mathfrak{D}$ mentioned by our theorem. Thus $\mathfrak{D}$ is not a circle, but it is convex and symmetric of order q, where

$$q \geq 2n + 1. \tag{4.5}$$

The domain $\mathfrak{D}$ lies in the z-plane. Let the origin be the center of symmetry of $\mathfrak{D}$.

Let the power series

$$z = c_0\zeta + c_1\zeta^{q+1} + c_2\zeta^{2q+1} + \cdots \tag{4.6}$$

express the conformal mapping of the unit circle in the ζ-plane onto $\mathfrak{D}$ in the z-plane; (4.6) maps the center of the circle into the center of symmetry of $\mathfrak{D}$. Since $\mathfrak{D}$ is convex, we can assert [3, p. 22, 161–163] that

$$|c_0| = r. \tag{4.7}$$

(c) We wish to apply the lemma of Section 3. We construct the trial functions $g_k(z)$ by *transplanting* the eigenfunctions of the unit circle into $\mathfrak{D}$. That is, we define $g_k(z)$ by postulating that

$$g_k(z) = \chi_k(\zeta) \tag{4.8}$$

when z and ζ are connected by (4.6). Since $\chi_k(\zeta)$ vanishes on the boundary of the unit circle, $g_k(z)$ must vanish on the boundary of $\mathfrak{D}$, and so it fulfills the condition (3.2) of the lemma.

We set $l(n) = l$ and, in view of (4.8), we set

$$\begin{cases} f(z) = x_1 g_1(z) + \cdots + x_l g_l(z) \\ \quad = x_1\chi_1(\zeta) + \cdots + x_l\chi_l(\zeta) = \varphi(\zeta). \end{cases} \tag{4.9}$$

We transform the integrals in the numerator and denominator of the RAYLEIGH ratio $R[f(z)]$, see (3.1), by introducing the polar coordinates ϱ, ϑ in the ζ-plane instead of the rectangular coordinates x, y in the z-plane; we have to use (4.9), (4.6), and (4.2) for this purpose. Observing that the numerator of the RAYLEIGH ratio (the DIRICHLET integral) is a conformal invariant [3, p. 95], we obtain

$$R[f(z)] = \frac{\iint |\operatorname{grad} \varphi(\zeta)|^2 \varrho \, d\vartheta \, d\varrho}{\iint [\varphi(\zeta)]^2 \left|\frac{dz}{d\zeta}\right|^2 \varrho \, d\vartheta \, d\varrho}. \tag{4.10}$$

It is understood that for these double integrals (and for the other double integrals that will be considered in the present section) the domain of integration is the unit circle, so that ϱ and ϑ vary from 0 to 1, and from $-\pi$ to π, respectively.

Set

$$\left|\frac{dz}{d\zeta}\right|^2 = \frac{dz}{d\zeta}\,\overline{\frac{dz}{d\zeta}} = \sum_{\nu=0}^{\infty} (\nu q + 1)^2 |c_\nu|^2 \varrho^{2\nu q} + H(\zeta). \tag{4.11}$$

According to (4.6), $H(\zeta)$ is an absolutely convergent double infinite series consisting of terms of the form

$$C \varrho^{h'} e^{ih\vartheta} \quad \text{with} \quad |h| \geq q \gtrless 2n + 1, \tag{4.12}$$

see (4.5); C is a constant, h' and h are integers, $h' > 0$. On the other hand, according to (4.9), the definition of $l = l(n)$, and (4.1), $[\varphi(\zeta)]^2$ is the sum of a finite number of terms which are all of the form

$$\psi(\varrho) e^{ik\vartheta} \quad \text{with} \quad |k| \leq 2n; \tag{4.13}$$

$\psi(\varrho)$ depends on ϱ alone (not on ϑ) and k is an integer. Integrating the product of the terms (4.12) and (4.13) with respect to ϑ from $-\pi$ to π, we get 0, and we conclude hence that

$$\iint [\varphi(\zeta)]^2 H(\zeta) \varrho \, d\vartheta \, d\varrho = 0. \tag{4.14}$$

It follows from (4.9), (4.10), (4.11), (4.14), and (4.7) that

$$\left\{ \begin{aligned} R[x_1 g_1 + x_2 + \cdots + x_l g_l] &= \frac{\iint |\operatorname{grad} \varphi(\zeta)|^2 \varrho \, d\vartheta \, d\varrho}{\iint [\varphi(\zeta)]^2 \sum_{\nu=0}^{\infty} (\nu q + 1)^2 |c_\nu|^2 \varrho^{2\nu q + 1} \, d\vartheta \, d\varrho} < \\ &< \frac{\iint |\operatorname{grad} \varphi(\zeta)|^2 \varrho \, d\vartheta \, d\varrho}{\dot{r}^2 \iint [\varphi(\zeta)]^2 \varrho \, d\vartheta \, d\varrho} \end{aligned} \right. \tag{4.15}$$

In fact, we can exclude the case of equality in the last step under (4.15), since $\mathfrak{D}$ is not a circle and so some of the coefficients $c_1, c_2, c_3, \ldots$ must be different from zero.

The last line of (4.15) is, by (4.9), the ratio of two quadratic forms in the variables $x_1, \ldots, x_l$. The roots of the corresponding characteristic equation are, as it is easily seen from Section 3 (b), $\lambda_1^{0\,2} \dot{r}^{-2}, \ldots, \lambda_l^{0\,2} \dot{r}^{-2}$ where $\lambda_1^0, \ldots, \lambda_l^0$

stand for the first l characteristic frequencies of the unit circle. It follows from (4.15), by virtue of the lemma of Section 3 and its corollary (a), that

$$\lambda_1 < \lambda_1^0 \dot{r}^{-1}, \ldots, \lambda_l < \lambda_l^0 \dot{r}^{-1}$$

which is the desired theorem.

5. *Additional remarks.* (a) *If the simply connected domain $\mathfrak{D}$ varies without any restriction, $\lambda_n \dot{r}$ possesses a finite upper bound that depends on n alone.*

This theorem, which I stated some time ago [*1*], can be proved by the method of Section 4. In fact, for the present purpose, we do not transplant all the successive eigenfunctions of the circle into $\mathfrak{D}$, as we did in Section 4 (c), but only those of type 0; we use the terminology of Section 4 (a). We obtain so easily the upper bounds, of which the first yields a "best" result [*3*, p. 97—98]:

$$\lambda_1 \dot{r} \leq 2.4048, \quad \lambda_2 \dot{r} \leq 5.5201, \quad \lambda_3 \dot{r} \leq 8.6537, \ldots^2).$$

The method of proof works also for the μ_n and yields the bounds:

$$\mu_2 \dot{r} \leq 3.8317, \quad \mu_3 \dot{r} \leq 7.0156, \quad \mu_4 \dot{r} \leq 10.1735, \ldots$$

Just the existence of a bound for $\mu_n \dot{r}$ follows from that for $\lambda_n \dot{r}$ since, as it is well known, $\mu_n < \lambda_n$.

(b) M. SCHIFFER found a very different method to prove the fact stated under (a). Indeed, if $\dot{r}$ is the maximum inner radius of the domain $\mathfrak{D}$, then $\mathfrak{D}$ contains, as is well known, a full circle of radius $\dot{r}/4$. Now the λ_n are monotonically decreasing domain functionals, and so each λ_n of $\mathfrak{D}$ is less than the corresponding λ_n of the circle with radius $\dot{r}/4$, which yields:

$$\lambda_1 \dot{r} \leq 9.619, \quad \lambda_2 \dot{r} \leq 15.328, \quad \lambda_3 \dot{r} \leq 15.328, \ldots$$

At the beginning, the bounds just obtained are higher than those obtained under (a), yet at the end, for sufficiently large n, the opposite is true.

(c) We do not know, for example, the least upper bound of $\lambda_2 \dot{r}$, or for which shape of $\mathfrak{D}$ this bound is attained or approximated. It may be worthwhile to mention that we can solve a similar problem.

Let A denote the area of $\mathfrak{D}$. Then

$$\lambda_2^2 A > 2j^2 \pi \tag{5.1}$$

where j, approximately 2.4048, is the first root of $J_0(x)$; the right hand side of (5.1) is the greatest lower bound of the left hand side.

This result is due to PETER SZEGÖ. It is based on the well known theorem, discovered by Lord RAYLEIGH and proved by G. FABER and E. KRAHN [*3*, p. VI, 191, 275],

$$\lambda_1^2 A \geq j^2 \pi. \tag{5.2}$$

[2]) Of course, the numerical values are only approximate and should be replaced by the appropriate zeros of the appropriate Bessel function.

We shall use (5.2) in its sharpest form: equality is attained only if $\mathfrak{D}$ is a circle.

We consider a more general situation: we assume that the nodal lines of the eigenfunction that corresponds to the characteristic frequency λ of $\mathfrak{D}$ dissect $\mathfrak{D}$ into s subdomains $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_s$ of area $A_1, A_2, \ldots, A_s$, respectively; of course, $s \geq 2$ and

$$A_1 + A_2 + \cdots + A_s = A. \tag{5.3}$$

The eigenfunction considered vanishes along the boundary of each of the subdomains $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_s$, and so λ is a characteristic frequency of each. Applying (5.2) s times, we obtain

$$\lambda^2 A_1 \geq j^2 \pi, \ldots, \lambda^2 A_s \geq j^2 \pi. \tag{5.4}$$

Equality cannot be valid in each of the s relations (5.4): otherwise each of the s subdomains would be an open circular disk which obviously could not be fitted together into one *connected* open domain $\mathfrak{D}$. Hence, adding the inequalities (5.4) and using (5.3) we obtain

$$\lambda^2 A > s j^2 \pi \tag{5.5}$$

which for $s = 2$ yields the desired (5.1). In fact, only the eigenfunction corresponding to λ_1 is without nodal lines, and so for λ_2 we must have $s \geq 2$. (Indeed, we have $s = 2$ for λ_2 by a theorem of COURANT.) This proof suggests also the shape of $\mathfrak{D}$ that renders the difference of the two sides of (5.1) infinitesimal: the union of two equal circles with given radius the common chord of which is of infinitesimal length.

Bibliography.

[1] PÓLYA, G.: Sur le rôle des domaines symétriques dans le calcul de certaines grandeurs physiques. Comptes Rendus Paris **235**, 1079—1081 (1952). — [2] PÓLYA, G., and M. SCHIFFER: Convexity of functionals by transplantation. J. Analyse Mathématique **3**, 245—345 (1954). — [3] PÓLYA, G., and G. SZEGÖ: Isoperimetric Inequalities. Princeton 1951. — [4] SZEGÖ, G.: Inequalities for certain eigenvalues of a membrane of given area. J. of Rational Mechanics and Analysis **3**, 343—356 (1954).

Department of Mathematics, Stanford University, Stanford, Cal. (U.S.A.).

(Eingegangen am 4. Juli 1955.)

LOEWNER, C.
Math. Zeitschr. Bd. 63, S. 338—340 (1955)

On totally positive matrices.

In memoriam ISSAI SCHUR.

By

CHARLES LOEWNER*).

A matrix $A = (a_{ij})$ is called totally positive if all its minors are non-negative. A product of two totally positive matrices is again totally positive. The totally positive square matrices of order n form therefore a semigroup S_n. The non-singular matrices of S_n represent then the elements of a subsemigroup s_n of the group of all real non-singular square matrices of order n[1]).

The semigroup s_n contains the identity matrix I. This fact allows the introduction of the concept of an infinitesimal element of s_n by the following definition adopted from the theory of Lie-groups: Let $U(t)$ be a differentiable matrix function of the real parameter t in an interval $0 \leq t \leq t_0$, representing for each t an element of s_n, and let $U(0) = I$. We call then the matrix $\left(\frac{dU(t)}{dt}\right)_{t=0}$ an infinitesimal element of s_n and the totality σ_n of all infinitesimal elements of s_n the infinitesimal semigroup of s_n.

Let now $\Omega(t)$ be a sectionwise continuous function of t $(0 \leq t \leq t_0)$ giving for each t an element of σ_n. The differential equation

$$\frac{dU(t)}{dt} = \Omega(t)\, U(t) \tag{1}$$

has a unique continuous solution satisfying the initial condition

$$U(0) = I. \tag{1'}$$

In the same way as in Lie-group theory one shows that $U(t_0)$ represents an element of s_n. It is called generated by the infinitessimal elements $\Omega(t)$ $(0 \leq t \leq t_0)$. In general, a semigroup cannot be completely recreated from its infinitesimal semigroup, even if the semigroup is connected. That this reconstruction is possible in case of our s_n is the essential content of this note. It may be considered as a grouptheoretical interpretation of a theorem by A. M. WHITNEY[2]).

*) This paper was written while the author was under contract with the Office of Scientific Research, Air Research and Development Command.

[1]) Regarding literature on totally positive matrices see e.g. I. J. SCHOENBERG: On smoothing operators and their generating function. Bull. Amer. Math. Soc. **50**, 199—230 (1953).

[2]) WHITNEY, A. M.: A reduction theorem for totally positive matrices. J. d'Analyse Math., Jerusalem **2**, 88—92 (1952).

We first determine the structure of σ_n by proving the

LEMMA. *A matrix $\Omega=(\omega_{ij})$ is an element of σ_n if and only if $\omega_{ij}=0$ for $|i-j|>1$ and all ω_{ij} $(|i-j|=1)$ are non-negative.*

PROOF. We use the above definition $\Omega=\left(\frac{dU(t)}{dt}\right)_{t=0}$ with $U(t)=(u_{ij}(t))$ representing for each t an element of s_n and $U(0)=I$. We have in particular $u_{ij}(t)\geq 0$ and $u_{ij}(0)=0$ for $i\neq j$. This leads to

$$\omega_{ij}\geq 0 \quad (i\neq j). \tag{2}$$

Consider next in case $n>2$ a second order minor

$$\begin{vmatrix} u_{kj}(t) & u_{kk}(t) \\ u_{ij}(t) & u_{ik}(t) \end{vmatrix} \qquad (i>k>j).$$

Since again it is non-negative and vanishes for $t=0$ its derivative at $t=0$ must be non-negative, i.e.

$$-\omega_{ij}\geq 0 \quad \text{for} \quad i-j>1. \tag{3'}$$

In the same way one proves that

$$-\omega_{ij}\geq 0 \quad \text{for} \quad j-i>1. \tag{3''}$$

By combining (3') and (3'') with (2) we see that the conditions of the lemma are necessary. In order to prove their sufficiency we start with the remark that σ_n forms a convex cone in the matrix space, i.e. that with two elements Ω_1 and Ω_2 of σ_n any linear combination $c_1\Omega_1+c_2\Omega_2$ with non-negative coefficients c_1 and c_2 belongs to σ_n. Indeed if $\Omega_1=\left(\frac{dU_1(t)}{dt}\right)_{t=0}$ and $\Omega_2=\left(\frac{dU_2(t)}{dt}\right)_{t=0}$, the matrices $U_1(t)$ and $U_2(t)$ all belonging to s_n and $U_1(0)=U_2(0)=I$, then $U(t)=U_1(c_1t)U_2(c_2t)$ also belongs to s_n and $U(0)=I$. Derivation at $t=0$ yields $\left(\frac{dU(t)}{dt}\right)_{t=0}=c_1\Omega_1+c_2\Omega_2$.

The preceding remark allows to consider only matrices Ω satisfying the condition of the lemma with only one non-zero element. If the latter is a diagonal element the matrices $e^{\Omega t}(t\geq 0)$ generated from the fixed Ω are diagonal matrices with positive diagonal elements and belong therefore to s_n. If the non-zero element is an $\omega_{ij}>0$ $(|i-j|=1)$ the matrices generated are

$$e^{\Omega t}=I+\omega_{ij}\,t\,E_{ij} \quad (t\geq 0), \tag{4}$$

E_{ij} being the matrix with all elements zero with the exception of a one at the crossing of the i-th row and the j-th column. But one verifies easily that the inequality $\omega_{ij}>0$ has the total positivity of the matrices (4) as consequence. This completes the proof of the lemma.

We now make use of the reduction theorem for totally positive matrices by A. M. WHITNEY[2]) which for non-singular square matrices of order n can be put into the form:

WHITNEY'S THEOREM. *Call* $I+\omega E_{ij}=F_{ij}(\omega)$. *Then every non-singular totally positive matrix* U *can be written as a product*

$$U = U_1 \cdot U_2 \dots U_{n-1} D V_1 \cdot V_2 \dots V_{n-1}, \tag{5}$$

where the U_i *and the* V_i, $i=1, 2, \dots, n-1$, *are of the form*

$$U_i = F_{n,n-1}(\omega^i_{n,n-1}) F_{n-1,n-2}(\omega^i_{n-1,n-2}) \cdots F_{i+1,i}(\omega^i_{i+1,i}) \tag{5'}$$

$$V_i = F_{n-i,n-i+1}(\omega^i_{n-i,n-i+1}) F_{n-i+1,n-i+2}(\omega^i_{n-i+1,n-i+2}) \cdots F_{n-1,n}(\omega^i_{n-1,n}) \tag{5''}$$

with all the ω*-s used non-negative, and* D *represents a diagonal matrix with positive diagonal elements.*

WHITNEY'S theorem gives a parametric representation of the totally positive matrices.

According to our lemma each factor in (5'), (5'') and D can be generated from an element of σ_n and we have the

THEOREM. *Each element of* s_n *can be generated from infinitesimal elements of* σ_n.

Department of Mathematics, Stanford University, Stanford, Cal. (U.S.A.).

(Eingegangen am 19. Juli 1955.)

MOTZKIN, TH. S.
Math. Zeitschr. Bd. 63, S. 341—344 (1955)

A proof of Hilbert's Nullstellensatz.

In memoriam Issai Schur.

By

THEODORE S. MOTZKIN.

Hilbert's Nullstellensatz (1891) was reduced by RABINOWITSCH (1929) to Theorem 1 below. For the reduction and further references see ZARISKI [3] who gives two ideal-theoretic proofs. The proof of BRAUER [1] for the unreduced theorem is quite elementary. The following short and straightforward proof presupposes very little field theory and uses no ideals. It also furnishes bounds[1]), albeit weak, on the degrees (Theorems 2 and 3) and on the number of polynomials needed and thence yields another incompatibility criterion (Theorem 5). Exact bounds for the degree in a particular case and for the number of polynomials are given in Theorems 4, 6 and 7 [2]).

Let K be a field, $x=(x_1, \ldots, x_n)$ an n-tuple of indeterminates, $K[x]$ the ring of polynomials and $K(x)$ the field of rational functions in x with coefficients in K. Let $\overline{K}$ be the algebraic closure of K.

Theorem 1. *If* $f_1, \ldots, f_r \in K[x]$ *are* $\neq 0, \ldots, 0$ *for every* $\xi=(\xi_1, \ldots, \xi_n)$, $\xi_1, \ldots, \xi_n \in \overline{K}$, *then* $\Sigma f_i g_i = 1$ *for some* $g_1, \ldots, g_r \in K[x]$.

Proof. $n=0$ is trivial; we use induction. Consider the f_i as polynomials in $K[x_2, \ldots, x_n]$ with coefficients in $K(x_1)$. By induction hypothesis[3]) we have $\Sigma f_i h_{i0} = a_0 \neq 0$, $h_{i0} \in K[x]$, $a_0 \in K[x_1]$. Let $a_0 = \Pi(x_1 - \xi_{1\lambda})$, $\xi_{1\lambda} \in \overline{K}$, and $K^* = K(\xi_{11}, \ldots, \xi_{1l})$. Again by induction hypothesis $\Sigma f_i(x_1=\xi_{1\lambda})\, \varphi_{i\lambda} = 1$, $\varphi_{i\lambda} \in K^*[x_2, \ldots, x_n]$. Hence $\Sigma f_i \varphi_{i\lambda} - 1 = (x_1 - \xi_{1\lambda})\psi_\lambda$, $\psi_\lambda \in K^*[x]$. Now let $a_1 = a_0/(x_1 - \xi_{11})$, and $h_{i1} = \varphi_{i1}\, a_1 - h_{i0}\psi_1$; then $\Sigma f_i h_{i1} = a_1$. Removing the factors of a_0 one by one we end up with $\Sigma f_i h_{il} = 1$, $h_{il} \in K^*[x]$. If K^* is an extension of K with basis $b_1 = 1, \ldots, b_\varrho, \ldots$, and if $h_{il} = \Sigma g_{i\varrho} b_\varrho$, $g_{i\varrho} \in K[x]$, then $\Sigma f_i g_{i1} = 1$, Q.E.D.

Theorem 2. *If all* f_i *in Theorem* 1 *are of degree* $\leq d$ *then the* g_i *may be chosen of degree* $\leq d_n \leq (d_{n-1}+d)\,\delta_{n-1} - d$, $d_0 = 0$, *where* δ_{n-1} *is defined in Theorem* 3 [4]).

[1]) The existence of a bound (maybe not independent of the number of polynomials) is stated, e.g., in VAN DER WAERDEN [2], p. 5, exercise.

[2]) This work was sponsored (in part) by the National Bureau of Standards and by the Offices of Naval Research and of Ordnance Research and presented to the American Mathematical Society on November 27, 1954.

[3]) Any $\eta_2, \ldots, \eta_n \in \overline{K(x_1)}$ annihilating all f_i would for some $\xi_1 \in \overline{K}$ have values $\xi_2, \ldots, \xi_n \in \overline{K}$ with the same property. This is a special case of [2], p. 51, first theorem. [The latter also implies, applied to the residue field (p. 53), the existence of a zero for a maximal ideal and thus, by the maximal condition, p. 22, for every ideal $\neq 1$: an immediate proof using some ideal theory and field theory.]

[4]) The polynomial 0 is considered as of degree $-\infty$.

The *proof*[5]) follows the above steps. By assumption we have deg $h_{i0} \leq \delta - d$, $l \leq \delta$, where $\delta = \delta_{n-1}$ and deg denotes the degree. By induction hypothesis deg $\varphi_{i\lambda} \leq d_{n-1}$ whence deg $\psi_\lambda \leq d_{n-1} + d - 1$ (note that for additional indeterminates as in Theorem 3 and $d > 1$ the degrees of the numerators of $\varphi_{i\lambda}$ and ψ_λ might be greater or undefined). Further deg $a_1 \leq \delta - 1$, deg $h_{i1} \leq 2d_{n-1} + \delta - 1$, deg $h_{i\lambda} \leq \delta - d + \lambda(d_{n-1} + d - 1)$, deg $g_{il} \leq$ deg $h_{il} \leq \delta (d_{n-1} + d) - d$, Q.E.D.

Let $y = (y_1, \ldots, y_m)$ be an m-tuple of additional indeterminates.

Theorem 3. *If* $f_1, \ldots, f_r \in K[x, y]$, *of degree* $\leq d$, *are* $\not\equiv 0, \ldots, 0$ *for every* $\xi = (\xi_1, \ldots, \xi_n)$, $\xi_1, \ldots, \xi_n \in \overline{K(y)}$, *then* $\Sigma f_i g_i = a \not\equiv 0$ *for some* $g_1, \ldots, g_r \in K[x, y]$ *of degree* $\leq \delta_n - d$ *and some* $a \in K[y]$ *of degree* $\leq \delta_n \leq \binom{n + d_n + d}{n+1} - \binom{n + d_n}{n+1}$, *where* d_n *is defined in Theorem* 2.

Note that the bounds given are not only independent of the individual f_i but also of r, m and K.

Proof. Let $f_i = \Sigma a_{ik} x^k$, $k = (k_1, \ldots, k_n)$, $0 \leq k_\nu$, $\Sigma k_\nu \leq d$, $x^k = \Pi x_\nu^{k_\nu}$, $a_{ik} \in K[y]$, deg $a_{ik} \leq d - \Sigma k_\nu$. By Theorem 2 we have $\Sigma f_i g_{i1} = 1$ for some $g_{i1} = \Sigma b_{ik'} x^{k'}$, $b_{ik'} \in K(y)$, $0 \leq k'_\nu$, $\Sigma k'_\nu \leq d_n$. Now $\Sigma f_i g_{i1} = 1$ means $\sum_{i,k} a_{ik} b_{ik'} = 0 (k'' = k + k'$ fixed $\not\equiv 0$, $0 \leq k''_\nu$, $\Sigma k''_\nu \leq d + d_n), = 1 (k'' = 0)$. These are $\binom{n + d_n + d}{n}$ linear equations in the $b_{ik'}$. Let the rank of their coefficient matrix be c and choose in it c independent columns. Then the c corresponding $b_{ik'}$ are, after setting all other $b_{ik'} = 0$, uniquely determined from some c equations which include the inhomogeneous one, as quotients of the minors of order $c - 1$ belonging to that equation by the c-order determinant $a \not\equiv 0$. Since the degree of the $(k''; i, k')$-element of a is $\leq d + \Sigma k'_\nu - \Sigma k''_\nu$ it follows that the degree of every term of a, as well as $d + \Sigma k'_\nu +$ the degree of every term of its $(0; i, k')$-minor, is $\leq \gamma = \Sigma(d + \Sigma k'_\nu - \Sigma k''_\nu)$ where γ is independent of the term selected. Choosing a non-vanishing term we see that deg a and $d +$ deg $a b_{ik'} x^{k'}$ are $\leq \gamma \leq \Sigma \min(d + d_n - \Sigma k''_\nu, d) = \Sigma(d + d_n - \Sigma k''_\nu) - \Sigma \max(0, d_n - \Sigma k''_\nu) \leq \binom{n + d_n + d}{n+1} - \binom{n + d_n}{n+1}$, Q.E.D.

For $d = 1$ Theorem 2 and 3 give $d_n = 0$, $\delta_n = 1$, the first of which expresses the duality theorem for linear equations. In this case the method of the proofs of Theorem 1 and 2 suffices to also obtain $\delta_n = 1$. For general d Theorem 2 and 3 imply $d_n \leq d^{e_n} - d$, $\delta_n \leq d^{n e_n + 1}$ (by $\delta_n \leq d\left(\binom{n + d_n + d}{n+1} - \binom{n + d_n + d - 1}{n+1}\right) = d\binom{n + d_n + d - 1}{n} \leq d(d_n + d)^n$), where $e_n = n!(1 + 1/1! + \cdots + 1/n!)$. Whereas perhaps $d_n \leq n d / \ln(n+1)$ and $d_n \leq n^{1 - 1/d}$ hold for $d \geq n$ and $d \leq n$, certainly $\delta_n \geq d^{n+1}$ as may be seen by considering general f_i for $r = n + 1$, $m = 1$. The bounds for d_n and δ_n drop somewhat by use of the following exact values for d_1 and δ_1.

Theorem 4. *In Theorem* 2 *and* 3, $d_1 = d - 1$ $(d > 0)$ *and* $\delta_1 = d^2$.

[5]) Simultaneous induction for Theorem 2 and 3.

Proof. For the proof of $d_1 = d-1$ we may suppose $a_{1d} \neq 0$. In $\Sigma f_i g_i = 1$ replace every g_i, $i>0$, by its remainder after division by f_1, while altering g_1 accordingly. This done, $\deg f_1 g_1 = \deg(1 - \sum_{i>1} f_i g_i) \leq 2d-1$, whence also $\deg g_1 \leq d-1$. Now $\delta_1 = d^2$ for $r=2$ by the reasoning in the proof of Theorem 3, with $\gamma = \Sigma d + \Sigma k' - \Sigma k'' = 2d^2 + 2\binom{d}{2} - \binom{2d}{2} = d^2$ (the matrix is the same as in the Sylvester resultant). For $r>2$ and infinite K, let $f_1^* = f_1$, $f_2^* = \Sigma z_i f_i$ with new indeterminates z_i and apply case $r=2$; then in $f_1^* g_1^* + f_2^* g_2^* = a^*$ replace the z_i by $\zeta_i \in K$ for which $a^*(\zeta_i) \neq 0$. To prove $\delta_1 = d^2$ for all K we use induction on r. Let be the greatest common divisor of $f_2, \ldots, f_r$ in $K[x, y]$ and hence also in $K(y)[x]$, and $d' = \deg f > 0$. Putting $f_i = \bar{f}_i f$, $i>1$, we have $\deg \bar{f}_i \leq d-d'$ and $\Sigma \bar{f}_i \bar{g}_i = \bar{a} \neq 0$ for some $\bar{g}_i \in K[x,y]$ of degree $\leq (d-d')^2 - (d-d')$ and some $\bar{a} \in K[y]$ of degree $\leq (d-d')^2$. On the other hand $f_1 \bar{g}_1 + fg = \bar{\bar{a}} \neq 0$ for some $\bar{g}_1$ and $g \in K[x, y]$, $\bar{\bar{a}} \in K[y]$, $\deg \bar{g}_1 \leq dd' - d$, $\deg g \leq dd' - d'$, $\deg \bar{\bar{a}} \leq dd'$; in fact $\gamma = dd'$ is seen to be the sum of the bounds for the degrees of the elements of the main diagonal of the Sylvester resultant of f_1 and f, after deducing $\deg_x \bar{g}_1 \leq d'-1$ and $\deg_x g \leq d-1$ as before. Hence $f_1 \bar{g}_1 \bar{a} + \Sigma f_i \bar{g}_i g = \bar{a}\bar{\bar{a}} \neq 0$ with $d + \deg \bar{g}_1 \bar{a}$, $d + \deg \bar{g}_i g$, $\deg \bar{a}\bar{\bar{a}}$ all $\leq dd' + (d-d')^2 \leq d^2$.

A polynomial h of degree $d+d'$ shall be called an *annihilator* of a polynomial f of degree d if $f(x_1^{-1}, \ldots, x_n^{-1})h$ contains no terms $\Pi x_\nu^{k_\nu}$, $k_\nu \geq 0$, $\Sigma k_\nu \leq d'$. If h has a non-vanishing constant term h is a *proper annihilator.*

Theorem 5. *The polynomials $f_1, \ldots, f_r \in K[x]$ of degree d have a common zero $\xi_1, \ldots, \xi_n \in \bar{K}$ if and only if they have a common proper annihilator $h \in K[x]$ of degree $d+d_n$, where d_n is defined in Theorem 2.*

Proof. The linear equations for b_{ik}, in the proof of Theorem 3 are, by duality, insolvable if and only if the equations $\sum_{k'' \geq k'} c_{k''} a_{i\,k''-k'} = 0$, $c_0 = 1$ for $c_{k''}$ are solvable. But this is tantamount to saying that $h = \Sigma c_{k''} x^{k''}$ is a proper annihilator of degree $d+d_n$ of every f_i.

The proof of Theorem 3 shows that no more than $\binom{n+d_n+d}{n}$ polynomials g_i are needed in Theorem 1 to 4. An improvement on this is obtained directly as follows.

Theorem 6. *In Theorem 1 to 4, all g_i but $\binom{n+d}{n}$ may be chosen as 0.*

Proof. The Theorem asserts that a smallest insolvable system of polynomial equations $f_i = 0$ has at most $r_0 = \binom{n+d}{n}$ members. Suppose it had $r > r_0$ members. Because of the minimality there exist points (n-tuples) $P_i = (\xi_{i1}, \ldots, \xi_{in})$, $\xi_{i\nu} \in \bar{K}$ such that $f_i(P_k) = 0$ for $i \neq k$. Since the f_i have r_0 coefficients one, say f_1, would be a linear combination of the others, whence $f_1(P_1) = 0$, against the insolvability supposition.

By the *Helly dimension* of a family of point sets we shall mean $r_0 - 1$, where r_0 is the smallest positive integer such that any subfamily with void

intersection must in turn contain a subfamily of $\leq r_0$ sets with void intersection. (One of the "Helly theorems" states that the family of compact convex sets in n-space has Helly dimension n.) Using the basis theorem ([2], p. 20), the proof of Theorem 6 implies:

Theorem 7. *The Helly dimension of the family of hypersurfaces of degree d in affine or projective n-space is $\leq$ its linear dimension $r_0 - 1 = \binom{n+d}{n} - 1$.*

In real or complex space, or any space over a field of infinite degree of transcendency, the hypersurfaces through $r_0 - 1$ from among r_0 general points will not have any point in common, whence r_0 points with this property can be found in every space over an infinite field, and the above Helly dimension is thus in these spaces exactly $r_0 - 1$.

References.

[1] BRAUER, R.: A note on Hilbert's Nullstellensatz. Bull. Amer. Math. Soc. **54**, 894—896 (1948). — [2] VAN DER WAERDEN, B. L.: Modern Algebra, vol. 2. New York 1950. — [3] ZARISKI, O.: A new proof of Hilbert's Nullstellensatz. Bull. Amer. Math. Soc. **53**, 362—368 (1947).

Added in proof. An ideal-theoretic proof of Theorem 1 is due to ANCOCHEA, G.: Sobre el "Nullstellensatz" de Hilbert. Ann. Mat. Pura Appl. (4) **29**, 31—34 (1949). A bound for d_n can be derived from HERMANN, G.: Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. Math. Ann. **95**, 736—788 (1926).

University of California, Los Angeles 24 (U.S.A.).

(Eingegangen am 26. Juli 1955).

BEHREND, F. A.
Math. Zeitschr. Bd. 63, S. 345—362 (1956)

A Contribution to the Theory of Magnitudes and the Foundations of Analysis

In memoriam ISSAI SCHUR

By

F. A. BEHREND

The main aim of this paper is to advocate a simultaneous development of the theory of magnitudes and the foundations of analysis.

Part I is an attempt to combine and simplify the axiomatic treatment of magnitudes and their measurement by real numbers as given by Hölder[1]), Bourbaki[2]), Zassenhaus[3]) and the author[4]). The main features of the system of axioms proposed in § 1.2 are: (1) The axioms apply to magnitudes, like lengths, with an unrestricted addition as well as to cases with restricted addition, e.g. to angles, or physical magnitudes with a restricted range of accessibility. (2) The inverse law and the positive law are assumed in the one-sided forms $(A_2^* r)$, $(AO_1 r)$ only. This permits the application of the axioms to irreversible magnitudes, like time intervals, which can only be laid off in one direction; in this case the equation $\alpha+\xi=\beta$ is solved by laying off α and β from the same initial point, but $\xi+\alpha=\beta$ cannot be solved similarly as α, β cannot be laid off from the same end-point; $\alpha<\alpha+\beta$ is a 'natural' law, $\beta<\alpha+\beta$ is not[5]). (3) The axioms are independent (§ 7). (4) The axioms guarantee the measurement of the magnitudes in the following precise sense: there exists an isomorphism (with respect to order and addition) of the magnitudes into the set of the positive real numbers; the map is uniquely determined by the choice of the unit (Theorem I, § 2).

However desirable and interesting, the theory of magnitudes, designed to elucidate the applicability of the real numbers, will be considered as a mere appendix to the foundations of analysis. But as is shown in Part II, the converse view is equally justified: the foundations of the real number system, including the exponential, logarithmic and trigonometric functions, may be regarded as a mere corollary to Theorem I (which therefore may well be styled the Fundamental Theorem of Analysis). The only background needed is a definition of the set $P^+(<, +)$ of the positive real numbers providing

[1]) HÖLDER, O.: Die Axiome der Quantität und die Lehre vom Maß. Ber. Verh. kgl.-sächs. Ges. Wiss., Leipzig, Math.-Phys. Kl. **53**, I, 1—64 (1901).

[2]) BOURBAKI, N.: Éléments de Mathématique, Livre III, Chap. V, § 2 (Mesure des grandeurs). Paris 1947.

[3]) ZASSENHAUS, H.: What is an angle? Amer. Math. Monthly **61**, 369—378 (1954).

[4]) BEHREND, F. A.: A system of independent axioms for magnitudes. Proc. Royal Soc. N.S.W. **87**, 27—30 (1953).

[5]) For an axiomatic treatment of time see loc. cit. [4]), 29—30.

the main properties of order and addition. Any of the traditional definitions may be used (Theorem I gives the isomorphism of all definitions). We choose the school representation of the positive real numbers as infinite decimal fractions, as the most elementary definition available: it may be obtained directly from the theory of the non-negative integers, without intermediate introduction of the positive rational numbers (§ 3.1–3). Inspite of being didactically the most natural method, this definition is usually rejected as being to unwieldy for the treatment of multiplication; but Theorem I supplies the automorphism group of $P^+(<, +)$ and hence the multiplication in P^+ (§ 3.5). The set $P^+(<, \cdot)$ is then the prototype, in multiplicative form, of the additive set $P(<, +)$ of all real numbers, and may be used to define P and to obtain the order and addition properties in P (§ 4.1–2). An extension of Theorem I (Theorem II, § 4.3) then furnishes the endomorphisms of P defining multiplication in P (§ 4.4), and the homomorphisms of $P(<, +)$ into $P^+(<, \cdot)$ defining the exponential and logarithmic functions (§ 5). The trigonometric functions are obtained by applying Theorem I to the system of angles (defined as rotations or unimodular complex numbers) and deducing the homomorphism of the group P on the group U of rotations (§ 6)[6].

This approach to the exponential, logarithmic and trigonometric functions is of course well-known in principle; but although it is carried out explicitly in Bourbaki[7], it seems to have received little attention[8], due, no doubt, to its being buried in too much topological background. The only topological tool used in the present exposition is Dedekind's section; the proof of Dedekind's law for the positive real numbers in decimal form takes only a few lines (§ 3.2); and the only essential use made of it is in the proof of Theorem I (§ 2.2); all the rest is pure algebra.

An Appendix contains proofs of the independence of a system of axioms for the positive real numbers (§ 7.1) and of the axioms of § 1.2 (§ 7.2), as well as a characterization, in terms of cyclic order, of the subgroups of the group U of rotations (§ 8).

I. Theory of Magnitudes

§ 1. Systems of magnitudes

1.1. Introduction

This section is concerned with (non-empty) sets $\Sigma = \{\alpha, \beta, \gamma, \ldots\}$ with

(i) *a relation* O: $\alpha < \beta$ (also written $\beta > \alpha$) is either true or false for any $\alpha, \beta \in \Sigma$,

[6]) For the sole purpose of developing the foundations of analysis, most of Part I is irrelevant as the commutative law in P^+ and U is trivial; all that is needed is (i) Theorem I for continua, (ii) Lemma I for the special case of U which can be obtained directly by representing the elements of U in the form $i^k \xi$ $(0 \leq k \leq 3)$ as in § 6.2 and then embedding U in the system of all formal expressions $i^n \xi$ on the lines of § 2.3.

[7]) loc. cit. [2]), Chap. V, § 4; VIII, § 2.

[8]) The response to several addresses on the subject, given in 1954, convinced me of this and encouraged the present publication in print.

(ii) *a (one-valued) operation A:* $\alpha+\beta$ is defined for some, not necessarily all, $\alpha, \beta \in \Sigma$; if $\alpha+\beta$ is always defined, Σ is called *complete*. From A we derive

(iii) *the operation N:* $n\alpha$ (n a natural number) is defined recursively, for some, not necessarily all, n by $1 \cdot \alpha = \alpha$, and $(n+1)\alpha = \alpha + n\alpha$ whenever $n\alpha$ and $\alpha + n\alpha$ are defined. If $n\alpha$ is not defined for all n, then n_α exists such that $n\alpha$ is defined for $n \leq n_\alpha$, undefined for $n > n_\alpha$. We write $\beta \dot{<} n\alpha$ if either $n\alpha$ is undefined or $n\alpha$ is defined and $\beta < n\alpha$.

1.2. Systems

$\Sigma = \Sigma(<, +)$ is called a *system* (of magnitudes) if it satisfies the following axioms:

(O_1) For any α, β exactly one of $\alpha < \beta$, $\alpha = \beta$, $\alpha > \beta$ holds *(trichotomy law)*,

($A_1 r$) $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$ if the right hand side is defined *(right associative law)*,

($A_2^* r$) If $\alpha < \beta$, then ξ exists such that $\alpha + \xi = \beta$ *(restricted right inverse law)*,

($AO_1 r$) $\alpha < \alpha + \beta$ *(right positive law)*,

(Arch) to any α, β there exists k such that $\beta \dot{<} k\alpha$ *(law of Archimedes)*.

1.3. Consequences of the axioms

($AO_2 r$) If $\alpha+\beta'$ is defined and $\beta < \beta'$, then $\alpha+\beta$ is defined and $\alpha+\beta < \alpha+\beta'$ *(right monotonic law)*.

Proof. $\alpha+\beta' = \alpha+(\beta+\xi) = (\alpha+\beta)+\xi > \alpha+\beta$ by ($A_2^* r$), ($A_1 r$), ($AO_1 r$).

($AO_2' r$) $\alpha+\beta < \alpha+\beta'$ implies $\beta < \beta'$, $\alpha+\beta = \alpha+\beta'$ implies $\beta = \beta'$ *(right cancellation law)*.

Proof. (O_1), ($AO_2 r$).

($NA_1 r$) $m\alpha + n\alpha = (m+n)\alpha$ if the right hand side is defined.

Proof. Induction with respect to m, using ($A_1 r$).

(NO_1) If $n\alpha$ is defined and $m < n$, then $m\alpha < n\alpha$.

Proof. ($NA_1 r$), ($AO_1 r$).

(NO_1') $m\alpha < n\alpha$ implies $m < n$.

Proof. (O_1), (NO_1).

($A_1 l$) $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$ if the left hand side is defined *(left associative law)*.

Proof. Put $\alpha+\beta = \sigma$, $(\alpha+\beta)+\gamma = \tau$, then $\alpha < \sigma$, $\sigma < \tau$ by ($AO_1 r$). ($A_1 l$) will now be proved if one of the following conditions (i) to (iv) is satisfied.

(i) $\alpha < \tau$, i.e. $\alpha + \xi = \tau$, by ($A_2^* r$). $\alpha+\beta < \alpha+\xi$ implies $\beta < \xi$ by ($AO_2' r$), $\beta+\eta = \xi$ by ($A_2^* r$), whence $(\alpha+\beta)+\gamma = \tau = \alpha+(\beta+\eta) = (\alpha+\beta)+\eta$ by ($A_1 r$) and $\eta = \gamma$ by ($AO_2' r$).

(ii) μ exists such that $\alpha < \mu$, $\sigma < \mu$, $\tau < \mu$.

Using (A_2^*r), $\mu = \tau + \xi = (\sigma + \gamma) + \xi = \sigma + (\gamma + \xi)$ by (i)
$= (\alpha + \beta) + (\gamma + \xi) = \alpha + (\beta + (\gamma + \xi))$ by (i)
$= \alpha + ((\beta + \gamma) + \xi) = (\alpha + (\beta + \gamma)) + \xi$ by (A_1r),

hence $\alpha + (\beta + \gamma)$ exists and is equal to τ by (A_1r).

(iii) Every ε which is the right hand term of a defined sum $\delta + \varepsilon$ satisfies $\varepsilon < \alpha$, $\varepsilon < \sigma$, $\varepsilon < \tau$.

γ is an ε, hence $\gamma < \alpha$, $\gamma + \alpha' = \alpha$, by (A_2^*r), and $\sigma = (\gamma + \alpha') + \beta = \gamma + (\alpha' + \beta)$ by (i) as $\gamma < \sigma$; hence

$\tau = \sigma + \gamma = (\gamma + (\alpha' + \beta)) + \gamma = \gamma + ((\alpha + \beta) + \gamma)$ by (i) as $\gamma < \tau$,
$= \gamma + (\alpha' + (\beta + \gamma))$, by (ii) as $\alpha', \alpha' + \beta, (\alpha' + \beta) + \gamma$ are ε's, hence $< \tau$,
$= (\gamma + \alpha') + (\beta + \gamma) = \alpha + (\beta + \gamma)$ by (A_1r).

(iv) There exists δ such that $n\delta$ is defined for all n. $n\delta + \eta$ is defined for all n, η (for $\eta < p\delta$ by (Arch) and $(n+p)\delta = n\delta + p\delta > n\delta + \eta$ by (NA_1r), (AO_2r)). It follows that $\xi < k\delta$ implies $\xi < (k+p)\delta$ (for $\xi = (k+p)\delta$ implies $\xi > k\delta$ by (NO_1), and $\xi > (k+p)\delta$ gives $(k+p)\delta + \eta = \xi$ by (A_2^*r), and $\xi = k\delta + (p\delta + \eta)$ by (A_1r), (NA_1r), as the latter sum is defined, and hence $\xi > k\delta$ by (AO_1r)). By (Arch), k, l, m exist such that $\alpha < k\delta$, $\sigma < l\delta$, $\tau < m\delta$, and putting $\mu = (k+l+m)\delta$, we have case (ii).

It will now be shown that at least one of conditions (i) to (iv) must hold. Suppose (i), (ii), (iii) are false. As (iii) is false, δ, ε exist such that $\delta + \varepsilon$ is defined and $\varepsilon \geqq \alpha$ or $\varepsilon \geqq \sigma$ or $\varepsilon \geqq \tau$ and, by (AO_2r), *one* of $\delta + \alpha$, $\delta + \sigma$, $\delta + \tau$ is defined. As $\alpha < \sigma$, $\sigma < \tau$ and $\tau \leqq \alpha$ [(i) being false], *all* of $\delta + \alpha$, $\delta + \sigma$, $\delta + \tau$ are defined. As (ii) is false, every ξ satisfies one of $\xi \leqq \alpha$, $\xi \leqq \sigma$, $\xi \leqq \tau$ and (AO_2r) implies that $\delta + \xi$ is defined for every ξ. In particular, $\delta + \delta = 2\delta$, $\delta + 2\delta = 3\delta$, etc., $\delta + (n-1)\delta = n\delta$ are all defined which gives case (iv).

(NA_1l) $m\alpha + n\alpha = (m+n)\alpha$ if the left hand side is defined.

Proof. (A_1l).

(N_1) $m(n\alpha) = (mn)\alpha$ if one side is defined.

Proof. (NA_1)[9].

(O_2) $\alpha < \beta$, $\beta < \gamma$ imply $\alpha < \gamma$ *(transitive law)*.

Proof. By (A_2^*r) $\alpha + \xi = \beta$, $\beta + \eta = \gamma$, hence $\gamma = (\alpha + \xi) + \eta = \alpha + (\xi + \eta) > \alpha$ by (A_1l), (AO_1r).

(Arch') If $\alpha < \beta$, then m exists such that $m\alpha < \beta \dot{\leqq} (m+1)\alpha$, whence $\beta = m\alpha + \varrho$, $\varrho \leqq \alpha$.

Proof. By (Arch) there exists a smallest m such that $\beta \dot{\leqq} (m+1)\alpha$, whence $m\alpha < \beta$, by (O_1), $m\alpha + \varrho = \xi$, by (A_2^*r), and $\varrho \leqq \alpha$ [as $\varrho > \alpha$ would imply $\beta = m\alpha + \varrho > m\alpha + \alpha = (m+1)\alpha$, by (AO_2r), (NA_1)].

(AO_1l) $\beta < \alpha + \beta$ *(left positive law)*.

[9]) Whenever the two forms (Lr), (Ll) of a law are valid, it will simply be quoted as (L)

Proof. Either: $\beta \leq \alpha$, then $\alpha < \alpha+\beta$, $\beta < \alpha+\beta$, by $(AO_1 r)$, (O_2); or: $\alpha < \beta$, then, by (Arch'), $m\alpha < \beta \leq (m+1)\alpha$ and $\alpha+\beta > \alpha+m\alpha = (m+1)\alpha \geq \beta$ using $(AO_2 r)$, (O_2).

$(AO_2 l)$ If $\alpha'+\beta$ is defined and $\alpha < \alpha'$, then $\alpha+\beta$ is defined and $\alpha+\beta < \alpha'+\beta$ *(left monotonic law)*.

Proof. $\alpha'+\beta = (\alpha+\xi)+\beta = \alpha+(\xi+\beta) > \alpha+\beta$, using $(A_2^* r)$, $(A_1 l)$, $(AO_1 l)$, $(AO_2 r)$.

(NO_2) If $n\beta$ is defined and $\alpha < \beta$, then $n\alpha$ is defined and $n\alpha < n\beta$.

Proof. (AO_2), (O_2).

(NO_2') $n\alpha < n\beta$ implies $\alpha < \beta$, $n\alpha = n\beta$ implies $\alpha = \beta$.

Proof. (NO_2), (O_1).

1.4. Discrete systems

A system Σ is called *discrete* if ν exists such that $\nu \leq \alpha$ for all α. Every α is of the form $n\nu$, for either $\alpha = 1 \cdot \nu$ or $\nu < \alpha$ whence, by (Arch'), $\alpha = (n-1)\nu + \varrho$, $\varrho \leq \nu$, i.e. $\varrho = \nu$, $\alpha = n\nu$. Σ thus consists exactly of all $n\nu$ with either $1 \leq n < \infty$ or $1 \leq n \leq n_\nu$.

1.5. Dense systems

A system is called *dense* if it is not discrete, i.e. if

(O_3) To any α there exists $\beta < \alpha$.

(O_3') To any α, k there exists β_k such that $k\beta_k < \alpha$.

Proof. β_1 exists by (O_3); if β_k exists then $k\beta_k + \xi = \alpha$ by $(A_2^* r)$, and by (O_3) there exists $\beta_{k+1} < \min(\beta_k, \xi)$ whence $\alpha = k\beta_k + \xi > k\beta_{k+1} + \beta_{k+1} = (k+1)\beta_{k+1}$, by (NO_2), (AO_2), (NA_1).

(O_3'') If $\gamma < \alpha$ then β exists such that $\gamma < \beta < \alpha$.

Proof. $\gamma + \xi = \alpha$, $\beta' < \xi$, $\gamma < \gamma + \beta' = \beta < \gamma + \xi = \alpha$, using $(A_2^* r)$, (O_3), (AO_1), (AO_2), (O_2).

1.6. The commutative law

Addition in every system is commutative:

(A_3) $\alpha+\beta = \beta+\alpha$ if one side is defined.

Proof. Let $\alpha+\beta$ be defined. By §1.4 and (NA_1), (A_3) holds in any discrete system. In a dense system distinguish two cases:

(i) $\alpha+\beta$ is not the greatest element of Σ, $\alpha+\beta < \gamma$, say, whence $(\alpha+\beta)+\xi = \gamma$, by $(A_2^* r)$. By (O_3') choose $2\delta < \min(\alpha, \beta, \xi)$, then, by (Arch'), $\alpha = m\delta + \varrho$, $\beta = n\delta + \sigma$, $\varrho \leq \delta$, $\sigma \leq \delta$, whence $\gamma = \alpha+\beta+\xi > m\delta + n\delta + 2\delta = (n+1)\delta + (m+1)\delta \geq \beta + \alpha$, by (AO_2), (NA_1), (O_2), so that $\beta+\alpha$ is defined. $\alpha+\beta < \beta+\alpha$ (and similarly $\beta+\alpha < \alpha+\beta$) leads to a contradiction choosing $\gamma = \beta+\alpha$ in the preceding argument. Hence $\alpha+\beta = \beta+\alpha$, by (O_1).

(ii) $\alpha+\beta$ is the greatest element of Σ. Choose $\delta<\alpha$, by (O_3), $\delta+\xi=\alpha$, by (A_2^*r), then $\alpha+\beta=\delta+\xi+\beta=\delta+\beta+\xi=\beta+\delta+\xi=\beta+\alpha$, applying (i) twice.

(NA_2) $n(\alpha+\beta)=n\alpha+n\beta$ if one side is defined.

Proof. (A_3).

1.7. Continuous systems

A dense system is called *continuous* if it satisfies

(Ded) If Σ is dissected into two non-empty classes Λ, Ω such that $\lambda\in\Lambda$, $\omega\in\Omega$ imply $\lambda<\omega$, then there exists σ (called a *generator* of the *section*) satisfying $\lambda\leq\sigma\leq\omega$ for all $\lambda\in\Lambda$, $\omega\in\Omega$ *(Dedekind's law)*. [The generator is unique by (O_3'').]

In any system satisfying (Ded), the axiom (Arch) may be replaced by the two axioms (A_1l), (AO_1l).

Proof. Assume (O_1), (A_1r), (A_1l), (A_2^*r), (AO_1r), (AO_1l) [and hence (O_2), (NO_1), (AO_2r)] and (Ded). Suppose (Arch) is false. Then there exist α, β such that $n\alpha$ is defined and $\leq\beta$ for all n. This implies $n\alpha<\beta$ [as $n\alpha=\beta$ implies $(n+1)\alpha>\beta$ by (NO_1)]. Define a section: $\lambda\in\Lambda$ if $\lambda\leq n\alpha$ for some n, $\omega\in\Omega$ if $\omega>n\alpha$ for all n. As $n\alpha\in\Lambda$ for all n and $\beta\in\Omega$, Λ, Ω are non-empty; and, by (O_2), $\lambda<\omega$. Let σ be a generator of the section; then $\alpha<2\alpha\leq\sigma$, $\alpha+\xi=\sigma$, by (A_2^*r), whence $\xi<\sigma$, by (AO_1l); this means $\xi\in\Lambda$, $\xi\leq n\alpha$ for some n, and $\sigma=\alpha+\xi\leq\alpha+n\alpha<(n+2)\alpha\leq\sigma$, using (AO_2r), (O_2), a contradiction.

A complete continuous system will be called a *continuum*. A continuum $\Gamma=\Gamma(<, +)$ is thus characterized by the axioms (O_1), (O_3), (Ded), (A_1), (A_2^*r), (AO_1r), (AO_1l), together with an unrestricted addition. In particular, the set $P^+(<, +)$ of the positive real numbers is a continuum (see § 3), which proves the *consistency* of the axioms; their *independence* will be proved in § 7.1; and it will follow from Theorem I (§ 2.23) that the axioms are *categorical* (in the sense that all continua are isomorphic).

§ 2. The measurement of magnitudes

2.1. The fundamental theorem

To measure the magnitudes of a system Σ means to set up an isomorphism of Σ into the set P^+ of the positive real numbers, i.e. a biunique map of Σ on a subset Σ' of P^+ which preserves order and addition (and hence also the operation N). A map $\alpha\to\alpha'$ of a system Σ into a system T is an isomorphism if and only if it is an A-homomorphism, i.e. $(\alpha+\beta)'=\alpha'+\beta'$. Order is automatically preserved, as $\alpha<\beta$ implies $\alpha+\xi=\beta$, $\alpha'+\xi'=\beta'$, $\alpha'<\beta'$; and this implies that the map is biunique. The existence of the required map, and its uniqueness (after choice of a unit) are expressed by

Theorem I. *If Σ is any system, Γ any continuum, and $\varepsilon\in\Sigma$, $\iota\in\Gamma$ any prescribed elements (units), then there exists exactly one isomorphism $\alpha\to\alpha'$ of Σ into Γ which maps ε on ι.*

Theorem I will first be proved for a complete system Σ (§ 2.2). The case of an incomplete Σ can be reduced to this case by embedding Σ in a complete system T (§ 2.3).

2.2. Proof of Theorem I for complete systems Σ

2.21. Existence of the isomorphism.

(i) Definition of the map. For any $\alpha \in \Sigma$ define α' as the generator of the following section in Γ:

$\lambda \in \Lambda_\alpha$ if m, n exist such that $m\varepsilon < n\alpha$ and $n\lambda < m\iota$,

$\omega \in \Omega_\alpha$ if $m\varepsilon < n\alpha$ always implies $n\omega \geq m\iota$.

Proof of the section properties. Λ_α is not empty: choose n, by (Arch), such that $\varepsilon < n\alpha$, and λ, by (O_3'), such that $n\lambda < \iota$, then $\lambda \in \Lambda_\alpha$. Ω_α is not empty: choose k such that $\alpha < k\varepsilon$, and let $\omega = k\iota$; then $m\varepsilon < n\alpha$ implies, together with $n\alpha < nk\varepsilon$, $m < nk$, whence $n\omega = nk\iota > m\iota$, i.e. $\omega \in \Omega_\alpha$. Finally, $\lambda \in \Lambda_\alpha$, $\omega \in \Omega_\alpha$ imply $n\lambda < m\iota \leq n\omega$, whence $\lambda < \omega$.

(ii) $(k\alpha)' = k\alpha'$.

Proof. If $k\alpha' < (k\alpha)'$, then $k\alpha' + \xi = (k\alpha)'$ and, choosing δ with $k\delta < \xi$, $k(\alpha' + \delta) < (k\alpha)'$ which means $m\varepsilon < nk\alpha$, $nk(\alpha' + \delta) < m\iota$ for suitable m, n, whence $\alpha' + \delta \in \Lambda_\alpha$, a contradiction.

If $(k\alpha)' < k\alpha'$, then $(k\alpha)' + \xi = k\alpha'$ and, with $k\delta < \xi$, $k(\alpha' - \delta) > (k\alpha)'$ which means that, for any m, n, $mk\varepsilon < nk\alpha$ always implies $nk(\alpha' - \delta) \geq mk\iota$ whence, cancelling k, $\alpha' - \delta \in \Omega_\alpha$, a contradiction.

(iii) $\alpha < \beta$ implies $\alpha' \leq \beta'$.

Proof. Λ_α is contained in Λ_β.

(iv) $(\alpha + \beta)' = \alpha' + \beta'$.

Proof. If Σ is discrete, this follows directly from (ii). Let Σ be dense, and suppose $(\alpha + \beta)' \neq \alpha' + \beta'$. Let the difference of the two terms be ξ. There exists δ such that $\delta < \alpha$, $\delta < \beta$, and $2\delta' < \xi$ [take any $\gamma \in \Sigma$, k such that $\gamma' < k\xi$, δ such that $2k\delta < \min(\alpha, \beta, \gamma)$, then by (ii), (iii), $2k\delta' \leq \gamma' < k\xi$, $2\delta' < \xi$]. By (Arch') m, n exist such that

$$m\delta < \alpha \leq (m+1)\delta,\ n\delta < \beta \leq (n+1)\delta, \text{ whence } (m+n)\delta < \alpha + \beta \leq (m+n+2)\delta.$$

By (ii), (iii) these inequalities imply

$$m\delta' \leq \alpha' \leq (m+1)\delta', \quad n\delta' \leq \beta' \leq (n+1)\delta',$$

and

$$(m+n)\delta' \leq (\alpha+\beta)' \leq (m+n+2)\delta';$$

but the first two of these also imply $(m+n)\delta' \leq \alpha' + \beta' \leq (m+n+2)\delta'$, so that $\xi \leq 2\delta'$, a contradiction.

(v) $\varepsilon' = \iota$.

Proof. As $m\varepsilon < n\varepsilon$ implies $m < n$ and $n\iota > m\iota$, $\iota \in \Omega_\varepsilon$, i.e. $\iota \geq \varepsilon'$. $\iota > \varepsilon'$ is impossible; for choose γ such that $\varepsilon' < \gamma < \iota$; then $m\varepsilon < (m+1)\varepsilon$ must imply, for all m, $(m+1)\gamma \geq m\iota$, i.e. $\gamma \geq m(\iota - \gamma)$, contradicting (Arch).

2.22. Uniqueness of the isomorphism.

Suppose there are two isomorphism $\alpha \to \alpha'$ and $\alpha \to \alpha^*$ with $\varepsilon' = \varepsilon^* = \iota$, and suppose $\alpha' \neq \alpha^*$ for some $\alpha, \alpha' < \alpha^*$, say, whence $\alpha' + \xi = \alpha^*$. Choose m such that $\iota < m\xi$, and let n be the smallest integer satisfying $m\alpha' < n\iota$; then $n\iota < m\alpha^*$ (as $n\iota \geq m\alpha^* = m\alpha' + m\xi > m\alpha' + \iota$ gives $(n-1)\iota > m\alpha'$). But these imply $m\alpha < n\varepsilon$ and $n\varepsilon < m\alpha$, a contradiction.

2.23. Corollary. *An isomorphism of a* continuum Σ *into the continuum* Γ *is an isomorphism of* Σ *on* Γ.

Proof. Suppose the isomorphism mapping ε on ι maps Σ on a proper subset of Γ. By Theorem I there exists an isomorphism of Γ into Σ mapping ι on ε. The resultant map is an isomorphism of Σ on a proper subset of Σ mapping ε on ε; but the identical map of Σ on Σ is the only such isomorphism.

2.3. Embedding of an incomplete system in a complete system

2.31. Lemma I. *Every incomplete system is a segment*[10] *of a complete system.*

Proof. By § 1.4 Lemma I is trivial for discrete systems. Let Σ be dense, $\sigma \in \Sigma$, and δ such that $3\delta < \sigma$. Let Δ be the set of all $\xi \leq \delta$; then (i) the sum of any two or three elements of Δ is defined in Σ, and (ii) every $\alpha \in \Sigma$ can be represented, by (Arch'), in the form $\alpha = k_\alpha \delta + \alpha^*$ where $\alpha^* \in \Delta$ and k_α is a non-negative integer (with the convention that a term $0 \cdot \delta$ is to be suppressed). Σ may thus be described as a set of expressions $x\delta + \xi$ where $\xi \in \Delta$, x a non-negative integer, and, obviously,

(E) $x\delta + \xi = y\delta + \eta$ if and only if $x = y$ and $\xi = \eta$,

(O) $x\delta + \xi < y\delta + \eta$ if and only if either $x < y$ or $x = y$ and $\xi < \eta$,

(A) $(x\delta + \xi) + (y\delta + \eta) = (x + y + k_{\xi+\eta})\,\delta + (\xi + \eta)^*$.

Let T be the set of *all* expressions $x\delta + \xi$ ($\xi \in \Delta$, $0 \leq x < \infty$). T becomes a complete system, and Σ a segment of T, on defining equality, order and addition in T by (E), (O), (A). The trichotomy law (O_1) holds by (E), (O), and the laws (A_1), $(A_2^* r)$, $(AO_1 r)$, (Arch) are proved as follows.

(A_1) $\big((x\delta + \xi) + (y\delta + \eta)\big) + (z\delta + \zeta)$

$$= (x + y + z + k_{\xi+\eta} + k_{(\xi+\eta)^*+\zeta})\,\delta + \big((\xi + \eta)^* + \zeta\big)^*$$

and $(x\delta + \xi) + \big((y\delta + \eta) + (z\delta + \zeta)\big)$

$$= (x + y + z + k_{\eta+\zeta} + k_{\xi+(\eta+\zeta)^*})\,\delta + \big(\xi + (\eta + \zeta)^*\big)^*$$

are equal; this is clear for $x = y = z = 0$ as $(\xi + \eta) + \zeta = \xi + (\eta + \zeta)$ in Σ, and hence holds for any x, y, z as the only term depending on x, y, z, viz. $x + y + z$, is common to both expressions.

$(A_2^* r)$ If $x\delta + \xi < z\delta + \zeta$ then $y\delta + \eta$ exists such that $(x\delta + \xi) + (y\delta + \eta) = z\delta + \zeta$, i.e. $x + y + k_{\xi+\eta} = z$, $(\xi + \eta)^* = \zeta$: if $\xi < \zeta$ put $\eta = \zeta - \xi$, $y = z - x$; if $\xi \geq \zeta$ (whence $x < z$) put $\eta = (\delta + \zeta) - \xi$, $y = z - x - 1$.

[10] A segment is the lower class of a Dedekind section.

(AO_1r) $x\delta+\xi<(x\delta+\xi)+(y\delta+\eta)$; for either $x<x+y+k_{\xi+\eta}$, or $y=k_{\xi+\eta}=0$ whence $\xi<\xi+\eta=(\xi+\eta)^*$.

(Arch) To any $x\delta+\xi$, $y\delta+\eta$ there exists n such that $y\delta+\eta<n(x\delta+\xi)$: if $x>0$ take $n=y+1$; if $x=0$ choose m such that $m\xi\leq\delta<(m+1)\xi$, then $(m+1)\xi\leq 2\delta$, whence $(m+1)(0\cdot\delta+\xi)=\delta+((m+1)\xi)^*$, and take $n=(m+1)(y+1)$.

2.32. Corollary. *Every incomplete continuous system is a segment of a continuum.*

Proof. (O_3) and (Ded) in Σ obviously imply (O_3) and (Ded) in T.

2.4. Proof of Theorem I for incomplete systems Σ

The existence of the isomorphism follows by embedding Σ in T, by § 2.31, and mapping T into Γ, by § 2.21. The uniqueness follows from § 2.22, observing that any isomorphism $\alpha\to\alpha'$ of Σ into Γ can obviously be extended to an isomorphism of T into Γ by mapping $x\delta+\xi$ on $x\delta'+\xi'$.

2.5. Remarks

The above embedding construction is closely related to Schreier's construction of a covering group[11]). A similar construction was proposed by Zassenhaus[12]): T can be defined as the set of all formal sums $\sum_{i=1}^{n}\xi_i$ with $\xi_i\in\Sigma$. This construction is more complicated than ours, chiefly because the definition of equality of formal sums is more involved than (E); on the other hand, Zassenhaus's construction can be applied to the more general case of non-archimedean systems.

Theorem I can also be proved directly without recourse to Lemma I [and even without using the commutative law (A_3)]. The proof is on the same lines as that given in § 2.2, but somewhat more laborious, as definition § 2.21 (i) cannot be used for systems with restricted addition and has to be replaced by: $\lambda\in\Lambda_\alpha$ if and only if m, n, δ exist such that $n\delta<\varepsilon$, $\alpha\dot{<}m\varepsilon$, and $n\lambda<m\iota$.

Lemma I and (A_3) can then be deduced from Theorem I and the existence of a continuum. Altogether, this procedure may be slightly shorter, but it is less transparent and rather artificial, as neither Lemma I nor (A_3) rest on the existence of a continuum.

Schreier's construction was used by Rieger[13]) to construct a linearly ordered covering group to any cyclically ordered group, and, in particular, to prove that the only group with continuous cyclic order is the group U of the unimodular complex numbers. It follows from Theorem I that the subgroups of U can be characterized as the groups with archimedean cyclic order (§ 8).

---

[11]) See e.g. H. Zassenhaus, The Theory of Groups, Chap. III, § 6. New York 1949.

[12]) loc. cit. [3]).

[13]) Rieger, L. S.: On the ordered and cyclically ordered groups, Věstnik Král. České Spol. Nauk, Tř. Mat.-Přir. **1946** no. 6; **1947** no. 1; **1948** no. 1.

II. Foundations of Analysis

§ 3. The positive real numbers

3.1. Definition

We assume only the theory of non-negative integers. A *positive real number* ξ is a sequence $x_0, x_1, x_2, \ldots, x_n, \ldots$ of non-negative integers, called *digits*, where $x_n \leq 9$ for $n > 0$ and $x_m > 0$ for infinitely many m. We write $\xi = x_0 x_1 x_2 \ldots x_n \ldots$, and if $x_m = 9$ for $m > n$, $\xi = x_0 x_1 x_2 \ldots x_n \dot{9}$. The set of all ξ will be called P^+.

3.2. Equality and order

Definitions. If $\xi = x_0 x_1 x_2 \ldots x_n \ldots$, $\eta = y_0 y_1 y_2 \ldots y_n \ldots$ then

(E) $\xi = \eta$ if and only if $x_n = y_n$ for all n;

(O) $\xi < \eta$ if and only if j exists such that $x_n = y_n$ for $n < j$, $x_j < y_j$.

Order satisfies the trichotomy law (O_1), the density law (O_3) (direct consequences of (E), (O)) and Dedekind's law (Ded).

Proof of (Ded). Let Λ, Ω be a section of P^+. A sequence of digits s_n satisfying

$$\omega_n = s_0 s_1 \ldots s_n \dot{9} \in \Omega \quad \text{and, if } s_n > 0, \quad \lambda_n = s_0 s_1 \ldots (s_n - 1)\, 9 \in \Lambda$$

can be constructed, recursively, as follows. As Ω is not empty, $u_0 u_1 \ldots u_n \ldots \in \Omega$, whence $u_0 \dot{9} \in \Omega$, and there exists a smallest integer s_0 such that $s_0 \dot{9} \in \Omega$. If s_m has been constructed for $m \leq n$, let s_{n+1} be the smallest integer such that $s_0 s_1 \ldots s_n s_{n+1} \dot{9} \in \Omega$. As Λ is not empty, not all numbers of the form $00 \ldots 0\dot{9}$ can be in Ω, i.e. not all s_n are zero. If $s_m \neq 0$ for infinitely many m, put $\sigma = s_0 s_1 \ldots s_n \ldots$; if $s_j \neq 0$, $s_m = 0$ for all $m > j$, put $\sigma = s_0 s_1 \ldots s_{j-1}(s_j - 1)\dot{9}$. If $\lambda \in \Lambda$ then $\lambda \leq \sigma$, as $\lambda > \sigma$ would imply $\lambda \geq \omega_n$ for some n; if $\omega \in \Omega$ then $\omega \geq \sigma$ as $\omega < \sigma$ would imply $\omega < \lambda_n$ for some n. Hence σ generates the section.

3.3. Addition

Definition. If ξ, η are as above, put

$e_n = 1$ if $m > n$ exists such that $x_i + y_i = 9$ for $n < i < m$, $x_m + y_m > 9$,

$e_n = 0$ otherwise,

$s_0 = x_0 + y_0 + e_0$ and $s_n = x_n + y_n + e_n - 10 e_{n-1}$ for $n > 0$.

Then $0 \leq s_n \leq 9$ for $n > 0$, as

either: $x_n + y_n < 9$ whence $e_{n-1} = 0$ and $0 \leq s_n = x_n + y_n + e_n \leq 9$,

or: $x_n + y_n = 9$ whence $e_{n-1} = e_n$ and $0 \leq s_n = 9(1 - e_n) \leq 9$,

or: $x_n + y_n > 9$ whence $e_{n-1} = 1$ and $0 \leq s_n = x_n + y_n + e_n - 10 \leq 9$;

and $s_m > 0$ for infinitely many m, as

either: $e_n = 0$ for $n \geq k$ whence $x_m > 0$, $s_m = x_m + y_m > 0$ for infinitely many $m > k$,

or: $e_n = 1$ for $n \geq k$ whence $x_m + y_m > 9$, $s_m = x_m + y_m - 9 > 0$ for infinitely many $m > k$,

or: $e_{m-1} = 0$, $e_m = 1$ for infinitely many m and $s_m = x_m + y_m + 1 > 0$.

Hence we may define

(A) $\xi+\eta=s_0 s_1 s_2 \dots s_n \dots$.

Addition is unrestricted and unique and satisfies the commutative law (A_3) (trivially), the positive law (AO_1), the associative law (A_1) and the restricted inverse law (A_2^*). We sketch the proofs of the last two:

Proof of (A_1). Put ${}^n\xi=\sum_{i=0}^{n} x_i 10^{n-i}$. It is easily seen that ${}^n(\xi+\eta)={}^n\xi+{}^n\eta+e_n$, whence ${}^n\xi+{}^n\eta\leq{}^n(\xi+\eta)\leq{}^n\xi+{}^n\eta+1$ and $|{}^n((\xi+\eta)+\zeta)-{}^n(\xi+(\eta+\zeta))|\leq 2$ which implies $(\xi+\eta)+\zeta=\xi+(\eta+\zeta)$ ($\sigma\neq\tau$ obviously implies that $|{}^n\sigma-{}^n\tau|$ is unbounded).

Proof of (A_2^*). Let $\xi<\zeta$; then $\xi=x_0 x_1 \dots x_{j-1} x_j x_{j+1} \dots$, $\zeta=x_0 x_1 \dots x_{j-1} z_j z_{j+1} \dots$ where $x_j<z_j$, and putting $\tau=x_0 x_1 \dots x_j \dot{9}$, $\xi\leq\tau<\zeta$. To solve $\xi+\eta=\zeta$, it is sufficient to solve $\tau+\sigma=\zeta$ and (if $\xi<\tau$) $\xi+\varrho=\tau$ [$\eta=\varrho+\sigma$ then satisfies $\xi+\eta=\zeta$ by (A_1)]. Obviously, $\sigma=00\dots0(z_j-x_j-1)\,z_{j+1}z_{j+2}\dots$; and $\varrho=00\dots0\,(9-x_{j+1})\,(9-x_{j+2})\dots$ unless ξ ends in a string of 9's; in this case $\xi=x_0x_1\dots x_j\dots x_k\,\dot{9}$, $x_k<9$, and $\varrho=00\dots0\,(9-x_{j+1})\dots(9-x_{k-1})\,(8-x_k)\,\dot{9}$.

3.4. Remarks

3.1–3.3 establish $P^+=P^+(<,+)$ as a continuum. — Instead of the base 10, any other base $b>1$ could have been used to define the positive real numbers; by Theorem I and the Corollary § 2.23 all these definitions are isomorphic (the isomorphism being uniquely determined by choosing $0\dot{9}$ resp. $0\,\overline{b-1}$ as units).

3.5. Multiplication

Definition. By Theorem I and the Corollary § 2.23 there exists exactly one automorphism α^* of $\Gamma=P^+$ (or any continuum) which maps a given (unit) element 1 ($1=0\dot{9}$, say) on any given α. The map $\alpha^*\to\alpha$ is thus a biunique map of the automorphism group G of Γ on Γ; it becomes an isomorphism of G on Γ by defining

(M) $\alpha\beta$ is the image of $\alpha^*\beta^*$.

This means that $\alpha\beta$ is the image of 1 by applying first β^*, then α^* to Γ; or, alternatively, that α^* is the map $\xi\to\alpha\xi$. 1^* is the identical map, i.e. the neutral element of G. — The laws of multiplication now follow easily:

(M_0) $1\cdot\alpha=\alpha\cdot1=\alpha$,

(M_1) $(\alpha\beta)\gamma=\alpha(\beta\gamma)$,

(M_2) For any α,β there exist ξ,η such that $\alpha\xi=\eta\alpha=\beta$,

express the group properties.

(MAr) $\alpha(\xi+\eta)=\alpha\xi+\alpha\eta$,

(MO_2r) $\xi<\eta$ implies $\alpha\xi<\alpha\eta$,

express the automorphism properties.

(MAl) $(\alpha+\beta)\xi=\alpha\xi+\beta\xi$.

Proof. The map $\xi \to \xi' = \alpha\xi + \beta\xi$ is an automorphism as $(\xi+\eta)' = \alpha(\xi+\eta) + \beta(\xi+\eta) = \alpha\xi + \alpha\eta + \beta\xi + \beta\eta = (\alpha\xi + \beta\xi) + (\alpha\eta + \beta\eta) = \xi' + \eta'$; as it maps 1 on $1' = \alpha \cdot 1 + \beta \cdot 1 = \alpha + \beta$, it must be identical with the automorphism $\xi \to (\alpha+\beta)\xi$.

(M_3) $\alpha\xi = \xi\alpha$.

Proof. The map $\xi \to \xi' = \xi\alpha$ is an automorphism as $(\xi+\eta)' = (\xi+\eta)\alpha = \xi\alpha + \eta\alpha = \xi' + \eta'$; as it maps 1 on $1' = 1 \cdot \alpha = \alpha$, it must be identical with the automorphism $\xi \to \alpha\xi$.

3.6. Structure of $\Gamma(<, \cdot)$

$\Gamma = \Gamma(<, \cdot)$ is a continuously ordered multiplicative group. Its structure may also be described as follows. The map $\alpha \to \bar{\alpha}$, where $\bar{\alpha}$ is the inverse of α, is an isomorphism of $\Gamma(<, \cdot)$ on $\Gamma(>, \cdot)$ [for $\alpha < \beta$ implies $\bar{\beta} = \bar{\beta}\alpha\bar{\alpha} < \bar{\beta}\beta\bar{\alpha} = \bar{\alpha}$, and $\overline{\alpha\beta} = \bar{\beta}\bar{\alpha} = \bar{\alpha}\bar{\beta}$, by (MO_2), (M_3)]. In particular, the set $\Gamma^+(<, \cdot)$ of all $\alpha > 1$ is mapped isomorphically on the set $\Gamma^-(>, \cdot)$ of all $\beta < 1$, and Γ^- consists of the inverses $\bar{\alpha}$ of the $\alpha \in \Gamma^+$. — $\Gamma^+(<, \cdot)$, and hence $\Gamma^-(>, \cdot)$, is a continuum.

Proof. The laws (O_1), (O_2), (Ded), (M_1) hold in Γ^+ as (O_1), (O_3''), (Ded), (M_1) hold in Γ. If $\alpha, \beta \in \Gamma^+$, i.e. $1 < \alpha$, $1 < \beta$, then $1 < \alpha = \alpha \cdot 1 < \alpha\beta$ by (MO_2) in Γ; this proves $\alpha\beta \in \Gamma^+$, i.e. the completeness of Γ^+, and $\alpha < \alpha\beta$ i.e. (MO_1) in Γ^+. If $\alpha < \beta$ then $\alpha\xi = \beta$ has a solution $\xi \in \Gamma$, and $\xi \in \Gamma^+$ (as $\xi \leqq 1$ would imply $\beta = \alpha\xi \leqq \alpha \cdot 1 = \alpha$); this is (M_3^*) in Γ^+. — $\Gamma(<, \cdot)$, and any set isomorphic with $\Gamma(<, \cdot)$, will be called a *bicontinuum*. In particular, $\Gamma(>, \cdot)$ is a bicontinuum.

§ 4. The real numbers

4.1. Definition

The set P of all real numbers is obtained by extending the set P^+ by the creation of a new element 0 and a new element $-\alpha$ to every $\alpha \in P^+$; P^- will denote the set of all $-\alpha$.

4.2. Order and addition

The definitions of order and addition are extended from $P^+(<, +)$ to P in the usual manner:

(O) $-\alpha < 0 < \beta$; $-\alpha < \beta$; $-\alpha < -\beta$ if and only if $\beta < \alpha$.

(A) $0 + \xi = \xi + 0 = \xi$ for all $\xi \in P$; $(-\alpha) + (-\beta) = -(\alpha + \beta)$;

$(-\alpha) + \beta = \beta + (-\alpha) = \beta - \alpha$ or 0 or $-(\alpha - \beta)$ if $\alpha <$ or $=$ or $> \beta$.

The laws of order and addition (O_1), (O_2), (O_3''), (Ded), (A_1), (A_2), (A_3), (AO_2) are obtained by proving that $P(<, +)$ is a bicontinuum, i.e. that $P(<, +)$ is isomorphic to $\Gamma(<, \cdot)$:

Lemma II. *There exists exactly one isomorphism $\xi \to \xi'$ of $P(<, +)$ on $\Gamma(<, \cdot)$ mapping a given $\varepsilon \in P^+$ on a given $\iota \in \Gamma^+$.*

Proof. By Theorem I and Corollary § 2.23, there exists an isomorphism $\alpha \to \alpha'$ of $P^+(<, +)$ on $\Gamma^+(<, \cdot)$ with $\varepsilon' = \iota$; this may be extended to an isomorphism of $P(<, +)$ on $\Gamma(<, \cdot)$ by defining $0' = 1$, $(-\alpha)' = \overline{\alpha'}$, as is easily verified using § 3.6, (O), (A). This is the only isomorphism with $\varepsilon' = \iota$; for the neutral element 0 must be mapped on the neutral element 1, hence the set P^+ of all $\alpha > 0$ must be mapped into the set Γ^+ of all $\beta > 1$, and $\alpha \to \alpha'$ is the only such isomorphism; as $(-\alpha) + \alpha = 0$ must imply $(-\alpha)'\alpha' = 1$, $(-\alpha)'$ must be $\overline{\alpha'}$.

4.3. Extension of Theorem I

Lemma II applies of course to *any* two bicontinua P, Γ. Observing that, by § 3.6, $P(>, +)$, $\Gamma(>, \cdot)$ are bicontinua, the Lemma may be extended to cover the case where ε, ι are *any* elements of P, Γ different from the neutral elements, by defining an *isomorphism* as a biunique map preserving the operation and *preserving or inverting* the order relation. It is convenient to include the case $\iota = 1$ as well. This may be done by defining a ***homomorphism*** as a map preserving the operation and preserving or inverting the relation $\leq$, and observing that there is only one homomorphism of P into Γ which is not an isomorphism of P on Γ, viz. the trivial homomorphism mapping all $\alpha \in P$ on 1. [If $\alpha \to \alpha'$ is a homomorphism which is not an isomorphism, then α, β exist such that $\alpha < \beta$, but $\alpha' = \beta'$; $\alpha + \delta = \beta$ then implies $\delta > 0$ and $\alpha'\delta' = \alpha'$, i.e. $\delta' = 1$ whence $(n\delta)' = 1$ for all n, and using (Arch), $\xi' = 1$ for all $\xi \in P$.] This proves

Theorem II. *If P, Γ are bicontinua, $\varepsilon \in P$, $\iota \in \Gamma$, and ε is different from the neutral element of P, then there exists exactly one homomorphism of P into Γ mapping ε on ι, and this is an isomorphism of P on Γ unless ι is the neutral element of Γ.*

4.4. Multiplication

Theorem II applies, in particular, to the endo- and automorphisms of $P(<, +)$. These define multiplication in P, and the laws of multiplication are obtained as in § 3.5, with the modifications caused by the existence of order-inverting automorphisms (multiplication by negative numbers) and the trivial endomorphism (multiplication by 0).

§ 5. Exponentials and logarithms

5.1. Definition

The image of ξ in the homomorphism $\xi \to \xi'$ of $P(<, +)$ on $\Gamma(<, \cdot) = P^+(<, \cdot)$ mapping 1 on ε is called ε^ξ, and the original of ξ' (when $\varepsilon \neq 1$) is called $\log_\varepsilon \xi'$.

5.2. Properties

The properties of ε^ξ (and similarly those of $\log_\varepsilon \xi'$) follow directly from the definition and the homomorphism properties:

(1) ε^ξ is defined for $\varepsilon > 0$ and all ξ; when ξ ranges over P, ε^ξ ranges over $\Gamma = P^+$ (except for $\varepsilon = 1$, when $\varepsilon^\xi = 1$).

(2) $\varepsilon^0 = 1,\ \varepsilon^1 = \varepsilon.$

(3) $\varepsilon^{\xi+\eta} = \varepsilon^\xi \varepsilon^\eta.$

(4) $\xi < \eta$ implies $\varepsilon^\xi <, =, > \varepsilon^\eta$ according to whether $\varepsilon >, =, < 1.$

The map $\xi \to \varepsilon^{\tau\xi}$ (τ fixed) is a homomorphism of P into Γ mapping 1 on ε^τ, hence identical with the homomorphism $\xi \to (\varepsilon^\tau)^\xi$, i.e.

(5) $\varepsilon^{\tau\xi} = (\varepsilon^\tau)^\xi.$

As ξ ranges over P, $\gamma = \varepsilon^\xi$ ranges over Γ, and as $\xi \to \tau\xi$ is an endomorphism of P, $\gamma = \varepsilon^\xi \to \varepsilon^{\tau\xi} = \gamma^\tau$ is an endomorphism of Γ; hence

(6) $(\gamma\delta)^\tau = \gamma^\tau \delta^\tau,$

(7) $\gamma < \delta$ implies $\gamma^\tau <, =, > \delta^\tau$ according to whether $\tau >, =, < 0.$

(8) The function $\varphi(\xi) = \varepsilon^\xi (\varepsilon \neq 1)$ is convex, i.e., if $\alpha < \beta$, and if $\lambda(\xi)$ is the linear function satisfying $\lambda(\alpha) = \varphi(\alpha)$, $\lambda(\beta) = \varphi(\beta)$, then $\varphi(\xi) < \lambda(\xi)$ for $\alpha < \xi < \beta$.

Proof. As ε^ξ is monotonic, by (4), and takes all real positive values, by (1), it is continuous; as is well known, the convexity then follows from $\varepsilon^{(\xi+\eta)/2} < (\varepsilon^\xi + \varepsilon^\eta)/2$ (for $\xi \neq \eta$) which is a consequence of (3). The appeal to continuity (extraneous to the methods used in this paper) is unnecessary; any monotonic $\varphi(\xi)$ satisfying $\varphi\big((\xi+\eta)/2\big) < \big(\varphi(\xi) + \varphi(\eta)\big)/2$ (for $\xi \neq \eta$) is convex[14]).

5.3. Remarks

The above properties are a sufficient basis for the calculus treatment of ε^ξ: as ε^ξ is convex it possesses, at every ξ, a right and a left hand differential coefficient and the two are equal because

$$\lim_{h \to +0} (\varepsilon^{\xi+h} - \varepsilon^\xi)/h = \lim_{h \to +0} \varepsilon^h (\varepsilon^{\xi-h} - \varepsilon^\xi)/(-h).$$

The use of limits may be avoided, if so desired. It is easy to show that a convex function $\varphi(\xi)$ possesses a linear support at every ξ_0, i.e. a linear function $\lambda(\xi) = \mu(\xi_0)(\xi - \xi_0) + \varphi(\xi_0) < \varphi(\xi)$ for $\xi \neq \xi_0$. If $\mu_\varepsilon \xi + 1$ is a linear support of ε^ξ at $\xi = 0$, then $\mu_\varepsilon \varepsilon^{\xi_0}(\xi - \xi_0) + \varepsilon^{\xi_0}$ is a linear support at ξ_0; and as the gradient $\mu_\varepsilon \varepsilon^{\xi_0}$ cannot belong to a support at any $\xi_1 \neq \xi_0$ and as, on the other hand, $\mu_\varepsilon \varepsilon^{\xi_0}$ ranges over all possible gradients, the support at every point is unique. As $\delta^\xi = (\varepsilon^{\log_\varepsilon \delta})^\xi = \varepsilon^{\xi \log_\varepsilon \delta}$, $\mu_\varepsilon \log_\varepsilon \delta \cdot \xi + 1 < \delta^\xi$ for $\xi \neq 0$, i.e. $\mu_\delta = \mu_\varepsilon \log_\varepsilon \delta$, whence $\varepsilon^{1/\mu_\varepsilon} = \delta^{1/\mu_\delta}$ for all δ. Putting $\varepsilon^{1/\mu_\varepsilon} = e$, we have $\mu_e = 1$, $\xi + 1 < e^\xi$ for $\xi \neq 0$, whence, for $\xi = 1/n$, resp. $-\big(1/(n+1)\big)$, $(1 + 1/n)^n < e < (1 + 1/n)^{n+1}$ and $\lim_{n \to \infty} (1 + 1/n)^n = e$.

[14]) The inequality implies $\varphi(\xi) < \lambda(\xi)$ for a dense set of ξ in (α, β); it follows that $\varphi(\gamma) \leq \lambda(\gamma)$ for every γ in (α, β), as $\varphi(\gamma) > \lambda(\gamma)$ would imply $\varphi(\xi) > \lambda(\xi)$ for a whole interval $\gamma \leq \xi < \delta$ (assuming φ increasing). Applying this result to the intervals $(\alpha, (\alpha+\beta)/2)$, $((\alpha+\beta)/2, \beta)$, it follows that $\varphi(\gamma) < \lambda(\gamma)$. — This is equivalent to a proof of the generalized theorem of the arithmetic and geometric mean: the convexity of ε^ξ means that $\varepsilon^{\tau\xi + (1-\tau)\eta} < \tau\,\varepsilon^\xi + (1-\tau)\,\varepsilon^\eta$ for $0 < \tau < 1$, $\xi \neq \eta$, whence, putting $\alpha = \varepsilon^\xi$, $\beta = \varepsilon^\eta$, $\alpha^\tau \beta^{1-\tau} < \tau\,\alpha + (1-\tau)\beta$.

§ 6. The trigonometric functions

6.1. Angles

Angles may be represented by the points $\mathfrak{x} = (\xi_1, \xi_2)$ on the unit circle $\xi_1^2 + \xi_2^2 = 1$ (i.e. by the unimodular complex numbers), and addition of angles may be described by means of rotations (or multiplication of complex numbers). Let U be the group of the unimodular complex numbers, i.e. the set of all pairs $\mathfrak{x} = (\xi_1, \xi_2)$ with $\xi_1^2 + \xi_2^2 = 1$, multiplication being defined by

$$\mathfrak{x}\mathfrak{y} = (\xi_1, \xi_2)(\eta_1, \eta_2) = (\xi_1\eta_1 - \xi_2\eta_2,\ \xi_1\eta_2 + \xi_2\eta_1).$$

6.2. Measurement of angles

U may be transformed into an incomplete system by means of a definition of order with $\mathfrak{e} = (1, 0)$ as the greatest element (four right angles) together with the corresponding restriction of the operation preventing the sum of two angles from exceeding $\mathfrak{e}$. It is simpler (though less elegant) to restrict the consideration to the set Σ of acute angles [including the right angle $\mathfrak{i} = (0, 1)$], i.e. the set of all $\mathfrak{x}$ with $1 > \xi_1 \geq 0$, $0 < \xi_2 \leq 1$. Every other element of U can then be represented uniquely in the form $\mathfrak{i}^k\mathfrak{x}$ with $1 \leq k \leq 3$, $\mathfrak{x} \in \Sigma$. Σ becomes an *incomplete continuous system* on defining

(O) $\mathfrak{x} < \mathfrak{y}$ if and only if $\xi_2 < \eta_2$ (i.e. $\xi_1 > \eta_1$),

(A) $\mathfrak{x} + \mathfrak{y} = \mathfrak{x}\mathfrak{y}$ if and only if $\mathfrak{x}\mathfrak{y} \in \Sigma$.

As the map $\mathfrak{x} \to \xi_2$ is an order isomorphism of Σ on the interval $0 < \xi_2 \leq 1$, order in Σ satisfies (O_1), (O_2), (O_3) and (Ded), and $\mathfrak{i}$ is the greatest element of Σ. Addition is commutative, and it remains to prove the following:

(AO_1) $\mathfrak{x} < \mathfrak{x} + \mathfrak{y}$.

Proof. $\xi_1 \geq \xi_1\eta_1 > \xi_1\eta_1 - \xi_2\eta_2$.

(A_2^*) If $\mathfrak{x} < \mathfrak{y}$ then $\mathfrak{z}$ exists such that $\mathfrak{x} + \mathfrak{z} = \mathfrak{y}$.

Proof. The solution $\mathfrak{z} = (\xi_1\eta_1 + \xi_2\eta_2,\ \xi_1\eta_2 - \xi_2\eta_1)$ of $\mathfrak{x}\mathfrak{z} = \mathfrak{y}$ lies in Σ, as $\xi_1\eta_1 + \xi_2\eta_2 > 0$ and $\xi_1\eta_2 - \xi_2\eta_1 > \eta_1\xi_2 - \xi_2\eta_1 = 0$.

(A_1) $(\mathfrak{x} + \mathfrak{y}) + \mathfrak{z} = \mathfrak{x} + (\mathfrak{y} + \mathfrak{z})$ if the left hand side is defined.

Proof. $\mathfrak{x} < \mathfrak{x} + \mathfrak{y}$, $\mathfrak{x} + \mathfrak{y} < (\mathfrak{x} + \mathfrak{y}) + \mathfrak{z}$, by (AO_1), hence, by (O_2), $\mathfrak{x} < (\mathfrak{x} + \mathfrak{y}) + \mathfrak{z} = (\mathfrak{x}\mathfrak{y})\mathfrak{z} = \mathfrak{x}(\mathfrak{y}\mathfrak{z}) = \mathfrak{x} + (\mathfrak{y} + \mathfrak{z})$ as $\mathfrak{y}\mathfrak{z} \in \Sigma$, by (A_2^*).

By Theorem I and § 2.32 there exists to every $\varrho > 0$ an isomorphism $\mathfrak{x} \to \tau$ of Σ into P^+ mapping $\mathfrak{i}$ on ϱ, and hence Σ on the segment Σ' of all τ with $0 < \tau \leq \varrho$. This provides a measure of the acute angles. (It may be extended to a map of U on $0 < \xi \leq 4\varrho$ by mapping $\mathfrak{i}^k\mathfrak{x}$ on $k\varrho + \tau$).

6.3. The trigonometric functions

Definition. Let $\mathfrak{x} = \mathfrak{t}(\tau)$ be the image of τ in the isomorphism of the interval $0 < \tau \leq \varrho$ on Σ. Then

(1) $\tau < \sigma$ implies $\mathfrak{t}(\tau) < \mathfrak{t}(\sigma)$,

(2) $\mathfrak{t}(\tau + \sigma) = \mathfrak{t}(\tau)\,\mathfrak{t}(\sigma)$ if $\tau + \sigma \leq \varrho$.

The definition of the function $\mathfrak{t}$ may be extended to all real arguments ξ by representing ξ in the form $\xi = k\varrho + \tau$ where k is an integer and $0 < \tau \leq \varrho$ and defining $\mathfrak{t}(\xi) = \mathfrak{i}^k\,\mathfrak{t}(\tau)$. Then

(2') $\mathfrak{t}(\xi+\eta) = \mathfrak{t}(\xi)\,\mathfrak{t}(\eta)$,

i.e. $\xi \to \mathfrak{t}(\xi)$ is a homomorphism of the additive group P on the multiplicative group U, and

(3) $\mathfrak{t}(\xi + 4\varrho) = \mathfrak{t}(\xi)$,

as $\mathfrak{i}^4 = \mathfrak{e}$ is the neutral element of U; hence U is isomorphic to the group, P mod 4ϱ. — If $\mathfrak{t}(\xi) = \mathfrak{x} = (\xi_1, \xi_2)$, define

$$\xi_1 = \cos_\varrho \xi, \quad \xi_2 = \sin_\varrho \xi.$$

Properties. $\cos_\varrho \xi$, $\sin_\varrho \xi$ are defined for all real ξ, have the period 4ϱ and satisfy the usual addition theorems, by (2'). They are determined by their behaviour in $0 \leq \xi \leq \varrho$; by (O) and (1) $\sin_\varrho \xi$ is monotonic increasing and $\cos_\varrho \xi$ monotonic decreasing in this interval; and as they assume every value from 0 to 1, they are continuous. If $0 \leq \xi < \eta \leq 2\varrho$, then $(\sin_\varrho \xi + \sin_\varrho \eta)/2 = \sin_\varrho\big((\xi+\eta)/2\big)\cos_\varrho\big((\xi-\eta)/2\big) < \sin_\varrho\big((\xi+\eta)/2\big)$ which implies the concavity of $\sin_\varrho \xi$ in $0 \leq \xi \leq 2\varrho$, as in § 5.2 (8) or [14]). It follows that $\sin_\varrho \xi$ possesses a right differential coefficient μ_ϱ at $\xi = 0$, and as $\sin_\varrho \xi$ is odd, this is equal to the left differential coefficient. The addition theorems then give $d \sin_\varrho \xi / d\xi = \mu_\varrho \cos_\varrho \xi$. As $\sin_\varrho(\varrho\xi) = \sin_\sigma(\sigma\xi)$, $\varrho\mu_\varrho = \sigma\mu_\sigma$ for all σ; putting $\varrho\mu_\varrho = \pi/2$, we get $\mu_{\pi/2} = 1$ and the functions $\sin \xi = \sin_{\pi/2}\xi$, $\cos \xi = \cos_{\pi/2}\xi$, with $d \sin \xi / d\xi = \cos \xi$, etc.

III. Appendix

§ 7. Independence proofs

7.1. Independence of the axioms for continua

For each of the axioms (O_1), (O_3), (Ded), (A_1), ($A_2^* r$), ($AO_1 r$), ($AO_1 l$) an example of a set $\Sigma(<, +)$ will be given satisfying all but this particular axiom. (O_1) will be split into two parts:

$\left(O_1 \begin{matrix} a \\ b \end{matrix}\right)$ For any α, β at $\begin{matrix} \text{most} \\ \text{least} \end{matrix}$ one of $\alpha < \beta$, $\alpha = \beta$, $\alpha > \beta$ holds.

Example $O_1 a$. $\Sigma = \{0\}$, defining $0 < 0$ and $0 + 0 = 0$.

Example $O_1 b$. $\Sigma = \{(\alpha_1, \alpha_2)\}$ where α_1, α_2 are any positive rational numbers, defining $(\alpha_1, \alpha_2) + (\beta_1, \beta_2) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2)$ and $(\alpha_1, \alpha_2) < (\beta_1, \beta_2)$ if and only if $\alpha_1 < \beta_1$ and $\alpha_2 < \beta_2$. [(Ded) holds vacuously as there are no Dedekind sections.]

Example O_3. Σ is the set of all natural numbers.

Example Ded. Σ is the set of all positive rational numbers.

Example A_1. Σ is the set of all positive real numbers with the usual order, addition $\dot{+}$ being defined by $\alpha \dot{+} \beta = \alpha + 2\beta$.

Example $A_2^* r$. Σ is the set of all real $\alpha > 1$.

Example AO_1r. Σ is the set of all continuous strictly increasing functions $f(\xi)$, defined for $0 \leq \xi \leq 1$, with $f(0) = \lambda_f > 0$, $f(1) = 1$. *Order:* The set of functions f with the same initial value $\lambda_f = \lambda$ has the cardinal $\aleph$, hence can be mapped biuniquely on the set $0 \leq \mu \leq 1$. Σ thus is in $1-1$-correspondence with the set (λ, μ) where $0 < \lambda < 1$, $0 \leq \mu \leq 1$, and may be ordered by ordering the (λ, μ) lexicographically. (O_1), (O_3), (Ded) then hold. *Addition:* $f \dot{+} g = g(f(\xi))$. (A_1) holds. (A_2r) holds, even without restriction: given any f, g, there exists x such that $f \dot{+} x = g$: for put $x(\xi) = g(f^{-1}(\xi))$ for $\lambda_f \leq \xi \leq 1$ (the inverse function f^{-1} of f is defined in this range) and $x(\xi) = (1 + \xi/\lambda_f)\lambda_g/2$ for $0 \leq \xi \leq \lambda_f$; then $x(f(\xi)) = g(f^{-1}(f(\xi))) = g(\xi)$. This shows that (AO_1r) must be false. (AO_1l) $g < f \dot{+} g$ is true, as $\lambda_{f+g} = g(f(0)) = g(\lambda_f) > g(0) = \lambda_g$. This shows, incidentally, that (A_2^*l) must be false as $y \dot{+} f = g$ implies $\lambda_f < \lambda_g$, hence cannot be solved when $f < g$ but $\lambda_f = \lambda_g$[15]).

Example AO_1l. $\Sigma = \{(n, \xi)\}$ where n is any non-negative integer, ξ any non-negative real number and $(n, \xi) \neq (0, 0)$. Order is lexicographic. Addition is defined by $(n, \xi) + (m, \eta) = (n + m, \eta)$ if $m > 0$ and $= (n, \xi + \eta)$ if $m = 0$.

7.2. Independence of the axioms § 1.2 for systems

Examples O_1a, O_1b, A_2^*r, of § 7.1 still aplly. Example AO_1l proves the independence of (Arch). For the independence of (AO_1r) see example 4 loc. cit. [4]). The independence of (A_1r) follows taking Σ as the set of all non-negative integers with the usual order and addition but letting $a + 0$ undefined for all a; (A_1r) is false as $0 + (0 + 1)$ is defined, but $(0 + 0) + 1$ is not. [(Arch) holds as 2.0 is undefined.] This example shows that (A_1r) cannot even be proved if (O_2), (A_1l) are valid.

§ 8. Groups with archimedean cyclic order

Let $G = \{\alpha, \beta, \gamma, \ldots\}$ be a (multiplicative) group with neutral element 1. G is a *group with archimedean cyclic order* if a triadic relation C, written (α, β, γ), is defined in G satisfying

(C_1) For any α, β, γ exactly one of the following three cases holds: (i) (α, β, γ), (ii) (α, γ, β), (iii) two or all of α, β, γ are equal *(trichotomy law)*,

(C_2) (α, β, δ), (β, γ, δ) imply (α, γ, δ) *(transitive law)*,

(MCr) (α, β, γ) implies $(\delta\alpha, \delta\beta, \delta\gamma)$ *(right monotonic law)*,

(ArchC) To any element α of infinite order there exists a natural number n such that $(\alpha^n, \alpha, 1)$ *(law of Archimedes)*.

Rieger's axiom[13])

(C_3) (α, β, γ) implies (β, γ, α)

[15]) It is an open question whether or not (AO_1l) can be replaced by (A_2^*l) in the system of axioms.

is a consequence of (C_1) and (C_2): for (α, β, γ) implies that α, β, γ are different: hence exactly one of (β, γ, α), (β, α, γ) is true; but the latter, together with (α, β, γ), would imply (α, α, γ) which is impossible. — The left monotonic law is not needed in the archimedean case.

It is easily verified that the group U of the unimodular complex numbers, and hence any subgroup of U, satisfies the above axioms if (α, β, γ) is defined to mean

$$i \begin{vmatrix} \alpha & \beta & \gamma \\ \alpha^{-1} & \beta^{-1} & \gamma^{-1} \\ 1 & 1 & 1 \end{vmatrix} > 0.$$

It will now be shown that, conversely, every group G with archimedean cyclic order is isomorphic to a subgroup of U. G may be transformed into a system $\Sigma(<, +)$ with greatest element 1 by defining

(O) $\alpha < \beta$ if and only if either $(\alpha, \beta, 1)$ or $\alpha \neq \beta = 1$,

(A) $\alpha + \beta = \alpha\beta$ if and only if $\alpha < \alpha\beta$.

(O_1) follows from (O) and (C_1), $(AO_1 r)$ from (A), and $(A_2^* r)$ from (A) and the right inverse law $(M_2 r)$ in G.

$(A_1 r)$ $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ if the right hand side is defined.

Proof. It is sufficient to show that the left hand side is defined (as both sides are then equal to $\alpha\beta\gamma$). As $\alpha + (\beta + \gamma)$ is defined, we have

(1) $(\alpha, \alpha\beta\gamma, 1)$ or $\alpha \neq \alpha\beta\gamma = 1$, whence $\beta\gamma \neq 1$,

(2) $(\beta, \beta\gamma, 1)$ [the alternative $\beta \neq \beta\gamma = 1$ is excluded by (1)].

From (2) and (MCr), $(\alpha\beta, \alpha\beta\gamma, \alpha)$ whence, together with (1), using (C_2) and (C_3),

(3) $(\alpha, \alpha\beta, 1)$, i.e. $\alpha < \alpha\beta$ so that $\alpha + \beta$ is defined,

(4) $(\alpha\beta, \alpha\beta\gamma, 1)$ or $\alpha\beta \neq \alpha\beta\gamma = 1$, i.e. $\alpha\beta < \alpha\beta\gamma$ so that $\alpha\beta + \gamma = (\alpha + \beta) + \gamma$ is defined.

(Arch) to any α, β there exists k such that $\beta < k\alpha$.

Proof. By (Arch$C$) $k > 1$ exists such that $(\alpha^k, \alpha, 1)$ or $\alpha^k = \alpha$, whence $\alpha^k \leq \alpha$, so that $\alpha + \alpha^{k-1}$, and hence $k\alpha$, are undefined.

By Lemma I and Theorem I (§ 2), Σ is isomorphic to a segment $0 < \xi \leq \varepsilon$ of a complete system T of positive real numbers, and hence G is isomorphic to the group $T \bmod \varepsilon$; but this is a subgroup of the group $P^+ \bmod \varepsilon$ which is isomorphic to U.

Department of Mathematics, University of Melbourne (Australia)

(Eingegangen am 19. Juli 1955)

SIEGEL, C. L.
Math. Zeitschr. Bd. 63, S. 363—373 (1956)

Die Funktionalgleichungen einiger Dirichletscher Reihen

ISSAI SCHUR zum Gedächtnis

Von

CARL LUDWIG SIEGEL

1. Die Eisensteinschen Reihen für die vierten bis achten Potenzen der Funktion

$$\vartheta(\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} \qquad (\tau = \xi + i\eta,\ \eta > 0)$$

sind zuerst von HARDY unter Benutzung der Kreismethode gewonnen worden. Bei Annäherung an einen rationalen Randpunkt a/c der oberen τ-Halbebene gilt nämlich

$$(c\tau - a)^{\frac{1}{2}}\,\vartheta(\tau) \to \gamma\left(\frac{a}{c}\right) \qquad \left(c > 0,\ (a, c) = 1\right),$$

wobei $\gamma(a/c)$ eine achte Einheitswurzel oder 0 ist, und als Eisensteinsche Entwicklung der Potenzen von $\vartheta(\tau)$ ergibt sich dann mittels der Theorie der Modulfunktionen

$$(1) \qquad \vartheta^m(\tau) = 1 + \sum_{a/c} \gamma^m\left(\frac{a}{c}\right)(c\tau - a)^{-\frac{m}{2}} \qquad (m = 4, \ldots, 8).$$

Hierin durchläuft a/c alle rationalen Zahlen, und im Falle $m = 4$ hat man wegen der bedingten Konvergenz zuerst über a bei festem c und sodann über c zu summieren. Die Hardysche Formel war der Ausgangspunkt für die Untersuchungen des Verfassers zur analytischen Theorie der quadratischen Formen, wodurch (1) als Spezialfall eines allgemeineren Satzes erkannt wurde.

Ein wichtiges Hilfsmittel in der Theorie der Eisensteinschen Reihen bildet bekanntlich der von Hecke eingeführte Konvergenz erzeugende Faktor $|c\tau - a|^{-z}$ $(z = x + iy,\ x > 0)$. Fügt man diesen Faktor dem allgemeinen Gliede in der Summe bei (1) hinzu, so entsteht die Dirichletsche Reihe

$$\Phi_m(\tau, z) = 1 + \sum_{a/c} \gamma^m\left(\frac{a}{c}\right)(c\tau - a)^{-\frac{m}{2}}\,|c\tau - a|^{-z},$$

welche in der Halbebene $x > 2 - \frac{m}{2}$ absolut konvergiert. Wie MAASZ erkannt hat, ist die analytische Funktion $\Phi_m(\tau, z)$ von z auch für $m = 1, 2, 3, 4$ nach $z = 0$ eindeutig fortsetzbar, und insbesondere ist $\Phi_m(\tau, 0) = \vartheta^m(\tau)$ für $m = 1, 3, 4$, während $\Phi_2(\tau, 0) = 2\vartheta^2(\tau)$ wird. Es ist nun bemerkenswert, daß sich $\Phi_m(\tau, z)$

als eine meromorphe Funktion von z erweist, welche einer einfachen Funktionalgleichung genügt. Dies ist übrigens nur ein Spezialfall eines Satzes über Dirichletsche Reihen, die sich in analoger Weise jeder ganzwertigen definiten oder indefiniten quadratischen Form von m Variabeln zuordnen lassen; doch soll weiterhin nur der Fall $m = 1$ diskutiert werden.

Verwendet man Riemanns Idee der Integraldarstellung von $\zeta(s)$ mittels der Thetafunktion, so gewinnt man aus $\Phi_1(\tau, z)$ eine meromorphe Funktion der zwei Veränderlichen s, z, welche für $z = 0$ in $\zeta(s)$ übergeht und sowohl in s als auch in z je einer Funktionalgleichung genügt. Es wäre noch von Interesse, die Verteilung der Nullstellen dieser Funktion in Abhängigkeit von dem Parameter z näher zu verfolgen.

2. Für die drei Funktionen

$$\vartheta_1(\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau}, \quad \vartheta_2(\tau) = \sum_{n=-\infty}^{\infty} (-1)^n e^{\pi i n^2 \tau}, \quad \vartheta_3(\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i (n+\frac{1}{2})^2 \tau}$$

gelten die Transformationsformeln

$$(2) \quad \begin{cases} \vartheta_1(\tau+1) = \vartheta_2(\tau), & \vartheta_2(\tau+1) = \vartheta_1(\tau), & \vartheta_3(\tau+1) = \sqrt{i}\,\vartheta_3(\tau), \\ \vartheta_1\left(-\frac{1}{\tau}\right) = \sqrt{\frac{\tau}{i}}\,\vartheta_1(\tau), & \vartheta_2\left(-\frac{1}{\tau}\right) = \sqrt{\frac{\tau}{i}}\,\vartheta_3(\tau), & \vartheta_3\left(-\frac{1}{\tau}\right) = \sqrt{\frac{\tau}{i}}\,\vartheta_2(\tau). \end{cases}$$

Um hieraus entsprechende Formeln für eine beliebige Modulsubstitution

$$\tau_M = \frac{a\tau + b}{c\tau + d}, \quad M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

zu gewinnen, beachte man, daß die Modulgruppe durch $\tau_1 = \tau + 1$ und $\tau_2 = -\tau^{-1}$ erzeugt wird. Es ergibt sich aus (2) ein Transformationsgesetz der Gestalt

$$(3) \quad \vartheta(\tau_M) = Q(M, \tau_M)\,\vartheta(\tau).$$

Dabei bedeutet $\vartheta(\tau)$ die aus $\vartheta_1(\tau), \vartheta_2(\tau), \vartheta_3(\tau)$ gebildete Spalte, und der Ausdruck

$$(4) \quad (c\tau + d)^{-\frac{1}{2}}\,Q(M, \tau_M) = (a - c\tau)^{\frac{1}{2}}\,Q(M, \tau)$$

ist eine von τ unabhängige monomiale dreireihige Matrix, deren von 0 verschiedene Elemente achte Einheitswurzeln sind. Für die Einheitsmatrix $M = E$ ist speziell $Q(E, \tau) = E$, und ebenso ist $Q(M, \tau) = Q(-M, \tau)$ trivial. Ist auch L eine Modulmatrix, so folgt aus (3) die Kompositionsformel

$$(5) \quad Q(ML, \tau_M) = Q(M, \tau_M)\,Q(L, \tau).$$

Zwecks näherer Untersuchung von $Q(M, \tau)$ als Funktion von M setze man in der Hermiteschen Bezeichnung

$$\Theta(\tau) = \Theta_{\delta\varepsilon}(\tau) = \sum_{n=-\infty}^{\infty} (-1)^{\varepsilon n} e^{\pi i \left(n + \frac{\delta}{2}\right)^2 \tau}$$

mit ganzen δ, ε und erhält in üblicher Weise aus der Zerlegung

$$\Theta\left(\tau+\frac{a}{c}\right)=\sum_{h=1}^{2c}(-1)^{\varepsilon h}\,e^{\pi i\left(h+\frac{\delta}{2}\right)^2\frac{a}{c}}\sum_{n=-\infty}^{\infty}e^{\pi i\left(2cn+h+\frac{\delta}{2}\right)^2\tau}\qquad(c>0)$$

bei dem Grenzübergang $\eta\to 0$, $\xi^2\eta^{-1}\to 0$ die Beziehung

$$(6)\qquad \sqrt{c\tau}\,\vartheta_k\left(\tau+\frac{a}{c}\right)\to\frac{1}{2}\sqrt{\frac{i}{c}}\sum_{h=1}^{2c}(-1)^{\varepsilon_k h}\,e^{\pi i\left(h+\frac{\delta_k}{2}\right)^2\frac{a}{c}}=\lambda_k\left(\frac{-a}{c}\right)\quad(k=1,2,3)$$

mit $\delta_1=0$, $\varepsilon_1=0$, $\delta_2=0$, $\varepsilon_2=1$, $\delta_3=1$, $\varepsilon_3=0$. Der Grenzwert $\lambda_k\left(\frac{-a}{c}\right)$ hängt also bei jedem festen c für $k=1,2$ nur von der Restklasse von a modulo $2c$ ab, und für $k=3$ nur von der Restklasse von a modulo $8c$. Ersetzt man in (6) noch τ durch

$$\tau_M-\frac{a}{c}=\frac{-1}{c(c\tau+d)}$$

mit $\xi=0$ und $\eta\to\infty$, so ergibt sich mittels (3), (4) die gewünschte Formel

$$(7)\qquad (c\tau-a)^{\frac{1}{2}}\,Q(M,\tau)\,\vartheta(\infty)=\lambda\left(\frac{-a}{c}\right)\qquad(c>0),$$

wobei $\vartheta_k(\infty)$ den Grenzwert von $\vartheta_k(\tau)$ für $\eta\to\infty$ bedeutet, nämlich $\vartheta_1(\infty)=1$, $\vartheta_2(\infty)=1$, $\vartheta_3(\infty)=0$, und $\lambda\left(\frac{-a}{c}\right)$ die Spalte der $\lambda_k\left(\frac{-a}{c}\right)$.

3. Man zerlege die Modulgruppe in Linksklassen bezüglich der Untergruppe der ganzen Modulsubstitutionen, bei denen also $c=0$ ist. Die Klassen sind dann den parabolischen Fixpunkten der Modulgruppe umkehrbar eindeutig zugeordnet, d.h. den rationalen Zahlen r mit Einschluß von ∞, und zwar liegen genau alle diejenigen $M=M_r$ in derselben Klasse, für welche $\frac{a}{c}=r$ ist. Indem man aus jeder Klasse einen Repräsentanten M_r herausgreift, bilde man mit der komplexen Variabeln $z=x+iy$ die Dirichletsche Reihe

$$(8)\qquad \varphi(\tau,z)=\eta^{z/2}\sum_r Q(M_r,\tau)\,\vartheta(\infty)\,|c\tau-a|^{-z}\qquad(x>\tfrac{3}{2}).$$

Vermöge (7) ist ersichtlich, daß die Glieder von der Auswahl der Klassenvertreter unabhängig sind. Bedeutet η_M den imaginären Teil von τ_M, so ist ferner

$$(9)\qquad \eta\,|c\tau-a|^{-2}=\eta_{M_r}^{-1}.$$

Nun durchläuft aber für irgendeine feste Modulmatrix M mit M_r auch MM_r ein volles System von Repräsentanten, und man kann in (8) daher M_r durch MM_r ersetzen. Aus (5), (9) folgt jetzt

$$(10)\qquad \varphi(\tau_M,z)=Q(M,\tau_M)\,\varphi(\tau,z),$$

das gleiche Gesetz der Transformation wie für $\vartheta(\tau)$.

Nach (7) gilt andererseits die Formel

$$(11)\quad \eta^{-\frac{z}{2}}\varphi(\tau,z)-\vartheta(\infty)=\sum_{a/c}\lambda\left(\frac{a}{c}\right)c^{-z-\frac{1}{2}}\left(\tau+\frac{a}{c}\right)^{-\frac{z+1}{2}}\left(\bar{\tau}+\frac{a}{c}\right)^{-\frac{z}{2}},$$

wo a/c alle gekürzten Brüche mit positiven Nennern durchläuft. Es ist die rechte Seite in eine Fouriersche Reihe zu entwickeln. Sind α, β komplexe Zahlen und die Realteile von $\alpha, \beta, \alpha+\beta-1$ positiv, so gilt die Fouriersche Entwicklung

$$i^{\alpha-\beta}\sum_{l=-\infty}^{\infty}(\tau+l)^{-\alpha}(\bar{\tau}+l)^{-\beta}=\sum_{n=-\infty}^{\infty}\varrho_n(\eta,\alpha,\beta)\,e^{2\pi i n\xi}$$

mit

$$(12)\quad \varrho_0(\eta,\alpha,\beta)=\frac{\Gamma(\alpha+\beta-1)}{\Gamma(\alpha)\,\Gamma(\beta)}(2\eta)^{1-\alpha-\beta},$$

$$(13)\quad (2\pi)^{-\alpha-\beta}\Gamma(\alpha)\,\Gamma(\beta)\,|n|^{1-\alpha-\beta}e^{2\pi|n|\eta}\varrho_n(\eta,\alpha,\beta)=\begin{cases}\sigma(4\pi n\eta,\alpha,\beta) & (n>0)\\ \sigma(-4\pi n\eta,\beta,\alpha) & (n<0),\end{cases}$$

$$(14)\quad \sigma(\eta,\alpha,\beta)=\int_0^{\infty}(u+1)^{\alpha-1}u^{\beta-1}e^{-\eta u}\,du.$$

Hierin ersetze man $\alpha, \beta, \tau, \eta$ durch $\frac{z+1}{2}, \frac{z}{2}, q_k^{-1}\left(\tau+\frac{a}{c}\right), q_k^{-1}\eta$ mit $q_1=q_2=2$, $q_3=8$. Entsprechend der oben bemerkten Periodizität von $\lambda_k\left(\frac{a}{c}\right)$ wird dann

$$(15)\quad \begin{cases}\displaystyle\sum_{a/c}\lambda_k\left(\frac{a}{c}\right)c^{-z-\frac{1}{2}}\left(\tau+\frac{a}{c}\right)^{-\frac{z+1}{2}}\left(\bar{\tau}+\frac{a}{c}\right)^{-\frac{z}{2}}\\ \qquad=\displaystyle\sum_{n=-\infty}^{\infty}j_k(n,z)\,\varrho_n\left(\frac{\eta}{q_k},\frac{z+1}{2},\frac{z}{2}\right)e^{2\pi i\frac{n}{q_k}\xi},\end{cases}$$

wobei die Abkürzung

$$j_k(n,z)=i^{-\frac{1}{2}}\sum_{a/c\,(\mathrm{mod}\,q_k)}\lambda_k\left(\frac{a}{c}\right)(q_kc)^{-z-\frac{1}{2}}e^{2\pi i\frac{na}{q_kc}}$$

eingeführt ist. Unter Benutzung der Definition von $\lambda_k\left(\frac{a}{c}\right)$ in (6) folgt

$$(16)\quad \zeta(z)\,j_k(n,z)=q_k^{-z-\frac{3}{2}}\sum_{c=1}^{\infty}c^{-z-1}\sum_{h\,(\mathrm{mod}\,q_kc)}(-1)^{q_kh}\sum_{a\,(\mathrm{mod}\,q_kc)}e^{2\pi i\frac{a}{q_kc}\left\{n-\frac{q_k}{2}\left(h+\frac{\delta_k}{2}\right)^2\right\}}.$$

4. Die innere Summe in (16) ist q_kc, falls

$$\frac{q_k}{2}\left(h+\frac{\delta_k}{2}\right)^2\equiv n\pmod{q_kc}$$

ist, und sonst 0. Daher ist insbesondere $j_3(n,z)=0$, wenn nicht $n\equiv 1\pmod 8$ ist, und

$$(17)\quad j_2(n,z)=(-1)^n j_1(n,z).$$

Weiterhin sei stets $n\equiv 1\pmod 8$ im Falle $k=3$.

Für jede natürliche Zahl m sei $A_m(n)$ die Anzahl der modulo m inkongruenten Lösungen von $x^2 \equiv n \pmod m$. Bedeutet p eine Primzahl, so setze man

$$f_p(n,z) = (1-p^{-z}) \sum_{l=0}^{\infty} A_{p^l}(n)\, p^{-lz} \qquad (p \neq 2),$$

$$f_2(n,z) = (2^z - 1) \sum_{l=1}^{\infty} A_{2^l}(n)\, 2^{-lz}$$

und erhält aus (16) die Produktdarstellungen

$$(18) \qquad j_1(n,z) = 2^{-z-\frac{1}{2}} \prod_p f_p(n,z), \quad j_3(n,z) = 2^{-2z+\frac{3}{2}} \prod_{p>2} f_p(n,z).$$

Es sei nun n^* die Diskriminante des durch $\sqrt{n}$ $(n \neq 0)$ erzeugten Körpers und $p^{2\alpha}$ die höchste in n/n^* aufgehende Potenz der Primzahl p, ferner $\chi_n(p) = \left(\frac{n^*}{p}\right)$ das Kroneckersche Symbol. Nach einfacher Rechnung ergibt sich dann

$$(19) \qquad \frac{1-\chi_n(p)\,p^{-z}}{1-p^{-2z}} f_p(n,z) = \begin{cases} p^{\alpha(1-2z)} + \left(1-\chi_n(p)\,p^{-z}\right) \sum_{l=0}^{\alpha-1} p^{l(1-2z)} & (p \neq 2) \\ \frac{1+\chi_n(2)}{1+2^{-z}} + \left(2^{1-z} - \chi_n(2)\right) \sum_{l=0}^{\alpha} 2^{l(1-2z)} & (p = 2). \end{cases}$$

Dies gilt sinngemäß auch für $n = 0$, nämlich

$$(1-p^{1-2z})\, f_p(0,z) = \begin{cases} 1-p^{-2z} & (p \neq 2) \\ (1-2^{-z})\,(1+2^{-z}) & (p = 2). \end{cases}$$

Daher wird

$$\prod_p f_p(0,z) = \frac{(1+2^{1-z})\,\zeta(2z-1)}{(1+2^{-z})\,\zeta(2z)}$$

und

$$(20) \qquad \prod_{p \nmid 2n} f_p(n,z) = \frac{L_n(z)}{\zeta(2z)} \prod_{p \mid 2n} \frac{1-\chi_n(p)\,p^{-z}}{1-p^{-2z}} \qquad (n \neq 0),$$

wo $L_n(z)$ die zum Charakter χ_n gehörige Dirichletsche L-Reihe bedeutet.

In Verbindung mit (17), (18), (19) erweist sich für $n \neq 0$ der Quotient von $(1+2^z)\,\zeta(2z)\, j_k(n,z)$ und $L_n(z)$ als eine ganze Funktion von z, nämlich als eine endliche Dirichletsche Reihe der Variabeln z, welche durch den Ausdruck

$$(21) \qquad \frac{3}{2} \sum_{t \mid \frac{4n}{n^*}} t^{\frac{1}{2}-x} = O\left(\left(\frac{4n}{n^*}\right)^{\mu}\right) \qquad \left(\mu > 0,\ \mu > \frac{1}{2} - x\right)$$

gleichmäßig in x majorisiert wird, während

$$(22) \qquad j_k(0,z) = \frac{(1+2^{1-z})\,\zeta(2z-1)}{2^{\frac{1}{2}}\,(1+2^z)\,\zeta(2z)} \qquad (k = 1, 2)$$

ist. Insbesondere ist $j_k(n,z)$ regulär in der Halbebene $x \geq \frac{1}{2}$, mit etwaiger Ausnahme von $z = 1$. In diesem Punkte tritt auch noch Regularität ein,

wenn nicht $n=0$ oder $n^*=1$ ist. In letzteren Fällen ist n eine Quadratzahl und $z=1$ ein Pol erster Ordnung. Als Residuum findet man $\pi^{-2}\sqrt{2}$ für $n=0$ und $\pi^{-2}\sqrt{8}$ $(k=1)$, $\pi^{-2}\sqrt{8}$ $(k=2)$, $\pi^{-2}\sqrt{2}$ $(k=3)$ für $n^*=1$.

5. Die analytische Fortsetzung von $\varrho_n\left(\frac{\eta}{q_k}, \frac{z+1}{2}, \frac{z}{2}\right)$ in die ganze z-Ebene ergibt sich zufolge (13), (14) aus der bekannten Darstellung der Funktion $\sigma(\eta, \alpha, \beta)$ durch ein Schleifenintegral. Danach ist $(1-e^{2\pi i\beta})\,\sigma(\eta, \alpha, \beta)$ eine ganze Funktion der Variabeln α, β, welche der Funktionalgleichung

$$\Gamma(1-\alpha)\,\sigma(\eta, \alpha, \beta) = \Gamma(\beta)\,\eta^{1-\alpha-\beta}\,\sigma(\eta, 1-\beta, 1-\alpha)$$

genügt. Daher wird

$$(23)\quad \begin{cases} \varrho_n\left(\frac{\eta}{q_k}, 1-\frac{z}{2}, \frac{1-z}{2}\right) \\ \qquad = \sqrt{\frac{2}{\pi}}\,\Gamma(z)\left(\frac{2\pi q_k|n|}{\eta}\right)^{\frac{1}{2}-z}\varrho_n\left(\frac{\eta}{q_k}, \frac{z+1}{2}, \frac{z}{2}\right)\begin{cases}\sin\frac{\pi z}{2} & (n>0)\\ \cos\frac{\pi z}{2} & (n<0),\end{cases} \end{cases}$$

und (12) liefert noch explizit

$$(24)\qquad \varrho_0\left(\frac{\eta}{2}, \frac{z+1}{2}, \frac{z}{2}\right) = \pi^{\frac{1}{2}}\,\frac{\Gamma(z-\frac{1}{2})}{\Gamma(z)}\,2^z\,\eta^{\frac{1}{2}-z}.$$

Benutzt man (11), (15) und die bei (21) gegebene Abschätzung, so wird ersichtlich, daß die drei Funktionen $\varphi_k(\tau, z)$ in z meromorph sind. Für $x>0$, $z\neq 1$ ist $\varphi_k(\tau, z)$ regulär. Bei $z=1$ liegt ein Pol erster Ordnung vor, und als Residuum ergibt sich

$$\pi^{-2}\sqrt{2}\left(2\pi + 2\sum_{m=1}^{\infty} 2\pi\, e^{\pi i m^2\tau}\right) = \pi^{-1}\sqrt{8}\,\vartheta_1(\tau) \qquad (k=1),$$

$$\pi^{-2}\sqrt{2}\sum_{m=0}^{\infty} 4\pi\, e^{\pi i(m+\frac{1}{2})^2\tau} = \pi^{-1}\sqrt{8}\,\vartheta_3(\tau) \qquad (k=3)$$

und analog $\pi^{-1}\sqrt{8}\,\vartheta_2(\tau)$ für $k=2$.

Man definiere noch

$$\pi^{-\frac{z}{2}}\,\Gamma\left(\frac{z}{2}\right)\zeta(z) = \varrho(z), \qquad \left(2^{\frac{z}{2}} + 2^{-\frac{z}{2}}\right)\varrho(2z) = \omega(z).$$

Dann ist die Funktion

$$(25)\qquad \varphi(\tau, z)\,\omega(z) = \psi(\tau, z)$$

ebenfalls meromorph in der Variabeln z, und in der Halbebene $x>0$ liegt nur ein Pol, und zwar von erster Ordnung, bei $z=1$. Das Residuum ist dort $\vartheta(\tau)$.

6. Es soll nun die Funktionalgleichung

$$(26)\qquad \psi(\tau, z) = \psi(\tau, 1-z)$$

bewiesen werden.

Mit der Abkürzung

$$\eta^{z/2}\,\omega(z)\,j_k(n,z)\,\varrho_n\Big(\frac{\eta}{q_k},\ \frac{z+1}{2},\ \frac{z}{2}\Big)=g_{kn}(\eta,z)\qquad(k=1,2,3)$$

gilt nach (11), (15), (22), (24), (25) die Fouriersche Entwicklung

$$\psi_k(\tau,z)=\vartheta_k(\infty)\Big(\eta^{\frac{z}{2}}\Big(2^{\frac{z}{2}}+2^{-\frac{z}{2}}\Big)\varrho(2z)+\eta^{\frac{1-z}{2}}\Big(2^{\frac{1-z}{2}}+2^{\frac{z-1}{2}}\Big)\varrho(2z-1)\Big)+$$
$$+\sum_{n\neq 0}g_{kn}(\eta,z)\,e^{2\pi i\frac{n}{q_k}\xi}.$$

Benutzt man jetzt (23) und die Funktionalgleichung von $\zeta(z)$, so folgt für die Differenz

$$\chi_k(\tau,z)=\psi_k(\tau,z)-\psi_k(\tau,1-z)$$

die Formel

$$\chi_k(\tau,z)=\sum_{n\neq 0}h_{kn}(\eta,z)\,e^{2\pi i\frac{n}{q_k}\xi}$$

mit

$$(27)\quad\begin{cases}h_{kn}(\eta,z)=g_{kn}(\eta,z)-g_{kn}(\eta,1-z)\\ \qquad=\eta^{z/2}\,\varrho_n\Big(\frac{\eta}{q_k},\ \frac{z+1}{2},\ \frac{z}{2}\Big)\Big(\omega(z)\,j_k(n,z)+\\ -\sqrt{\frac{2}{\pi}}\,\Gamma(z)\,(2\pi q_k|n|)^{\frac{1}{2}-z}\,\omega(1-z)\,j_k(n,1-z)\begin{cases}\sin\frac{\pi z}{2}\\ \cos\frac{\pi z}{2}\end{cases}\end{cases}\qquad\begin{matrix}(n>0)\\ (n<0).\end{matrix}$$

Man beachte, daß der Ausdruck in der großen Klammer nicht von η abhängt.

Da $\varrho_n\Big(\frac{\eta}{q_k},\ \frac{z+1}{2},\ \frac{z}{2}\Big)$ zufolge (13), (14) für $\eta\to\infty$ exponentiell verschwindet, so ist $\eta^{\frac{1}{4}}\,\chi_k(\tau,z)$ als Funktion von τ im Fundamentalbereich der Modulgruppe beschränkt, also nach (10), (25) in der ganzen oberen τ-Halbebene. Daher ist auch

$$(28)\qquad h_{kn}(\eta,z)=q_k^{-1}\int_0^{q_k}\chi_k(\tau,z)\,e^{-2\pi i\frac{n}{q_k}\xi}\,d\xi=O(\eta^{-\frac{1}{4}})\qquad(\eta\to 0).$$

Andererseits gelten nach (13), (14) für $\eta\to 0$ die Beziehungen

$$\sigma(\eta,\alpha,\beta)\sim\Gamma(\alpha+\beta-1)\,\eta^{1-\alpha-\beta},$$

$$(29)\qquad\varrho_n\Big(\frac{\eta}{q_k},\ \frac{z+1}{2},\ \frac{z}{2}\Big)\sim\sqrt{2\pi}\,\frac{\Gamma(z-\frac{1}{2})}{\Gamma(z)}\Big(\frac{\eta}{q_k}\Big)^{\frac{1}{2}-z}\qquad(x>\tfrac{1}{2}),$$

wobei die Realteile von $\alpha+\beta-1,\beta$ als positiv vorausgesetzt werden. Für $x>\frac{3}{2}$ ist nun aber

$$\eta^{-\frac{1}{4}}=o\Big(\eta^{\frac{1-x}{2}}\Big)\qquad(\eta\to 0).$$

Aus (27), (28), (29) folgt demnach

$$(30)\quad \omega(z)\,j_k(n,z) = \sqrt{\frac{2}{\pi}}\,\Gamma(z)\,(2\pi q_k|n|)^{\frac{1}{2}-z}\,\omega(1-z)\,j_k(n,1-z)\begin{cases}\sin\frac{\pi z}{2} & (n>0)\\ \cos\frac{\pi z}{2} & (n<0)\end{cases}$$

und

$$h_{kn}(\eta,z)=0,\qquad \chi_k(\tau,z)=0,$$

also die Behauptung (26).

In (30) sind wegen (18), (20) die Funktionalgleichungen der L-Reihen $L_n(z)$ enthalten. Umgekehrt läßt sich aus diesen wiederum (30) und damit (26) folgern.

Da $\omega(z)$ bei $z=0$ einen Pol erster Ordnung vom Residuum -1 hat, so erhält man

$$(31)\qquad \varphi(\tau,0)=\vartheta(\tau).$$

7. Es verdient bemerkt zu werden, daß sich (26) auch ohne Benutzung der Funktionalgleichung für $\zeta(z)$ gewinnen läßt, wobei sich dann diese zugleich ergibt. Für die Funktion

$$\mu_k(\tau,z)=\varrho(1-2z)\,\psi_k(\tau,z)-\varrho(2z)\,\psi_k(\tau,1-z)$$

findet man die Fouriersche Entwicklung

$$\mu_k(\tau,z)=\vartheta_k(\infty)\,\eta^{\frac{1-z}{2}}\left(2^{\frac{1-z}{2}}+2^{\frac{z-1}{2}}\right)\times$$
$$\times\left(\varrho(1-2z)\,\varrho(2z-1)-\varrho(2z)\,\varrho(2-2z)\right)+\sum_{n\neq 0} f_{kn}(\eta,z)\,e^{2\pi i\frac{n}{q_k}\xi},$$

wo $f_{kn}(\eta,z)$ aus dem Ausdruck für $h_{kn}(\eta,z)$ in (27) dadurch hervorgeht, daß man bei $\omega(z),\omega(1-z)$ die Faktoren $\varrho(1-2z)$, $\varrho(2z)$ hinzufügt. Für $x>\frac{3}{2}$ gilt dann auch (28) mit $f_{kn}(\eta,z)$, $\mu_k(\tau,z)$ an Stelle von $h_{kn}(\eta,z)$, $\chi_k(\tau,z)$, und man kann die frühere Schlußweise durchführen. Es folgt jetzt

$$f_{kn}(\eta,z)=0,\qquad \varrho(1-2z)\,\varrho(2z-1)=\varrho(2z)\,\varrho(2-2z),\qquad \mu_k(\tau,z)=0.$$

Nach (18), (19) ist speziell

$$j_1(1,z)=\frac{1+2^{-z}+2^{1-2z}}{2^{\frac{1}{2}}(1+2^z)}\,\frac{\zeta(z)}{\zeta(2z)},$$

und aus dem Verschwinden von $f_{11}(\eta,z)$ erhält man weiter

$$\varrho(1-2z)\,\varrho(z)=\varrho(2z)\,\varrho(1-z),$$

also auch

$$\frac{\varrho(z)}{\varrho(1-z)}=\frac{\varrho(2^{-m}z)}{\varrho(1-2^{-m}z)}\qquad (m=1,2,\ldots).$$

Mit den leicht direkt zu beweisenden Formeln

$$\zeta(z)\sim\frac{1}{z-1}\qquad (z\to 1),\qquad \zeta\left(\frac{1}{2}\right)\neq 0$$

folgt dann

$$\varrho(z)\sim\frac{1}{z-1},\qquad \varrho(1-z)\sim\frac{1}{z-1}\qquad (z\to 1).$$

Der Grenzübergang $m \to \infty$ ergibt

$$\varrho(z) = \varrho(1-z)$$

und (26).

8. Anstatt das asymptotische Verhalten von $\sigma(\eta, \alpha, \beta)$ für $\eta \to 0$ heranzuziehen, kann man für den Nachweis des identischen Verschwindens von $\chi_k(\tau, z)$ auch folgendermaßen eine Differentialgleichung benutzen.

Sind α, β, γ Konstanten, so genügt die Funktion

$$w = \eta^\alpha \tau^\beta \bar{\tau}^\gamma$$

der homogenen linearen elliptischen Differentialgleichung

$$\eta^2 (w_{\xi\xi} + w_{\eta\eta}) + \eta (i\mu w_\xi - \nu w_\eta) + \lambda w = 0 \tag{32}$$

mit

$$\mu = \beta - \gamma, \quad \nu = 2\alpha + \beta + \gamma, \quad \lambda = \alpha(1 + \alpha + \beta + \gamma) = \alpha(1 - \alpha + \nu).$$

Diese Differentialgleichung bleibt ungeändert, wenn τ durch $\tau + h$ mit einer reellen Konstanten h ersetzt wird. Sie geht ferner in sich über, wenn α, β, γ durch

$$\alpha' = 1 + \alpha + \beta + \gamma, \quad \beta' = -1 - \gamma, \quad \gamma' = -1 - \beta$$

ersetzt werden.

Es sei u der reelle Teil einer Lösung $w = u + iv$ von (32), und es seien λ, μ, ν reell. Nach (32) ist dann

$$\eta^2 (u_{\xi\xi} + u_{\eta\eta}) - \eta (\mu v_\xi + \nu u_\eta) + \lambda u = 0.$$

Hat nun $w\bar{w} = u^2 + v^2$ irgendwo ein Maximum > 0 und ist dort w reell, also $v = 0$, $u^2 > 0$, so ist dort

$$u_\xi = 0, \quad u_\eta = 0, \quad u u_{\xi\xi} + v_\xi^2 \leq 0, \quad u u_{\eta\eta} + v_\eta^2 \leq 0$$

$$u(u_{\xi\xi} + u_{\eta\eta}) + v_\xi^2 + v_\eta^2 \leq 0$$

$$(2\eta v_\xi + \mu u)^2 + (2\eta v_\eta)^2 \leq (4\lambda + \mu^2) u^2$$

$$4\lambda + \mu^2 \geq 0. \tag{33}$$

Speziell sei $\alpha = \frac{z}{2} + \frac{1}{4}$, $\beta = -\frac{z}{2} - \frac{1}{2}$, $\gamma = -\frac{z}{2}$, also $\mu = -\frac{1}{2}$, $\nu = 0$, $4\lambda = 4\alpha(1-\alpha) = 1 - \left(z - \frac{1}{2}\right)^2$, ferner $\alpha' = \frac{3}{4} - \frac{z}{2}$, $\beta' = \frac{z}{2} - 1$, $\gamma' = \frac{z}{2} - \frac{1}{2}$. Da $\nu = 0$ ist, so genügt $w = \eta^\alpha$ ebenfalls der Differentialgleichung (32), also nach (11) auch die Funktion $\eta^{\frac{1}{4}} \varphi(\tau, z)$; zunächst für $x > \frac{3}{2}$, und sodann nach dem Prinzip der analytischen Fortsetzung in der ganzen z-Ebene. Nun gehen aber α', β', γ' aus α, β, γ durch Vertauschung von z mit $1 - z$ hervor, und folglich ist auch die Funktion $\eta^{\frac{1}{4}} \varphi(\tau, 1-z)$ eine Lösung von (32). Damit ist auch

$$\delta(\tau, z) = \eta^{\frac{1}{4}} \mu(\tau, z) = \eta^{\frac{1}{4}} \left(\varrho(1-2z)\, \omega(z)\, \varphi(\tau, z) - \varrho(2z)\, \omega(1-z)\, \varphi(\tau, 1-z)\right)$$

als Lösung von (32) erkannt.

Andererseits ist nach (10) der Ausdruck

$$\sum_{k=1}^{3} \delta_k \bar{\delta}_k = \sigma = \sigma(\tau, z)$$

in Abhängigkeit von τ bei der Modulgruppe invariant. Speziell sei z reell und $z = x > \frac{3}{2}$. Dann strebt $\sigma(\tau, z)$ gegen 0, wenn τ im Fundamentalbereich der Modulgruppe nach ∞ geht. Ist σ nicht identisch 0, so wird für ein endliches τ ein positives Maximum erreicht. Indem man noch $\delta_1, \delta_2, \delta_3$ einer geeigneten konstanten unitären Transformation unterwirft, kann man voraussetzen, daß bei jenem τ die Werte $\delta_1 = \sigma^{\frac{1}{2}}$, $\delta_2 = 0$, $\delta_3 = 0$ sind. Dann hat also $\delta_1 \bar{\delta}_1$ selber bei τ ein positives Maximum, und δ_1 ist dort reell. Da für reelles z auch α, β, γ reell sind, so liefert (33) die Bedingung

$$\left(x - \frac{1}{2}\right)^2 \leq \frac{5}{4}, \qquad x \leq \frac{1+\sqrt{5}}{2}.$$

Für $x > \frac{1}{2}(1+\sqrt{5})$ folgt also $\sigma = 0$, womit wieder die Funktionalgleichung bewiesen ist.

9. Nach dem Vorbild von Riemanns Integraldarstellung der Zetafunktion bilde man mit $s = \sigma + it$ die Funktion

$$\varrho(s,z) = \frac{1}{2} \int_0^\infty \eta^{\frac{s}{2}} \left(\varphi_1(i\eta, z) - \eta^{\frac{z}{2}} - \frac{\omega(1-z)}{\omega(z)} \eta^{\frac{1-z}{2}} \right) \frac{d\eta}{\eta}, \tag{34}$$

wobei über positive η integriert wird. Aus den Reihendarstellungen (11), (15) findet man in Verbindung mit der Abschätzung bei (21), daß das Integral in der Halbebene $\sigma > \mathrm{Max}(1+x, 2-x)$ absolut konvergiert und gliedweise Integration vorgenommen werden kann. So erhält man in dieser Halbebene die Entwicklung

$$\varrho(s,z) = 2^{z-1} \pi^{\frac{z-s}{2}} \frac{\Gamma\left(\frac{s+z}{2}\right)}{\Gamma(z)} \sum_{n=1}^{\infty} n^{\frac{z-s-1}{2}} \left(j_1(n,z)\, G(s,z) + j_1(-n,z)\, H(s,z) \right) \tag{35}$$

mit

$$G(s,z) = \int_0^\infty u^{\frac{z}{2}-1} (u+1)^{-\frac{s+z}{2}} (u+2)^{\frac{z-1}{2}}\, du,$$

$$H(s,z) = \int_0^\infty u^{\frac{z-1}{2}} (u+1)^{-\frac{s+z}{2}} (u+2)^{\frac{z}{2}-1}\, du.$$

Andererseits ist

$$\varphi_1(i\eta^{-1}, z) = \eta^{\frac{1}{2}} \varphi_1(i\eta, z),$$

und dies ergibt die für alle s, z gültige Formel

$$\left\{ \begin{aligned} \varrho(s,z) = {} & \frac{1}{2} \int_1^\infty \left(\eta^{\frac{s}{2}} + \eta^{\frac{1-s}{2}} \right) \varphi_1(i\eta, z) \frac{d\eta}{\eta} + \\ & - \left(\frac{1}{s+z} + \frac{1}{1-s+z} \right) - \frac{\omega(1-z)}{\omega(z)} \left(\frac{1}{1+s-z} + \frac{1}{2-s-z} \right). \end{aligned} \right. \tag{36}$$

Folglich ist $z(1-z)(s+z)(1-s+z)(1+s-z)(2-s-z)\,\omega(z)\,\varrho(s,z)$ eine ganze Funktion von s, z, und es gelten die beiden Funktionalgleichungen

$$\varrho(1-s,z)=\varrho(s,z), \qquad \omega(z)\,\varrho(s,z)=\omega(1-z)\,\varrho(s,1-z).$$

Wegen (31) geht (36) bei $z=0$ in

$$\varrho(s,0)=\frac{1}{2}\int_1^\infty\left(\eta^{\frac{s}{2}}+\eta^{\frac{1-s}{2}}\right)\vartheta_1(i\eta)\,\frac{d\eta}{\eta}-\frac{1}{s(1-s)}$$

über, während (35) zu

$$\varrho(s,0)=\sum_{n=1}^{\infty}n^{-s}=\zeta(s)$$

führt.

Trägt man in (34) für $\varphi_1(i\eta,z)$ nach (11) die Eisensteinsche Reihe ein, so gewinnt man schließlich die Entwicklung

$$2\varrho(s,z)=\sum_{\substack{a\\ c>0}}a^{\frac{s-z-1}{2}}\,c^{-\frac{s+z}{2}}\left(\lambda_1\left(\frac{a}{c}\right)P(s,z)+\lambda_1\left(\frac{-a}{c}\right)Q(s,z)\right)$$

mit

$$P(s,z)=\int_0^\infty\eta^{\frac{s-z}{2}-1}(1+i\eta)^{-\frac{z+1}{2}}(1-i\eta)^{-\frac{z}{2}}\,d\eta,$$

$$Q(s,z)=\int_0^\infty\eta^{\frac{s+z}{2}-1}(1+i\eta)^{-\frac{z}{2}}(1-i\eta)^{-\frac{z+1}{2}}\,d\eta$$

im Streifen $2-x<\sigma<x-1$. Für $x>\frac{3}{2}$ läßt sich diese Formel benutzen, um die Verteilung der Nullstellen von $\varrho(s,z)$ auf $\sigma=\frac{1}{2}$ zu untersuchen.

Göttingen, Mathematisches Institut der Universität

(Eingegangen am 27. Juli 1955)

FREUDENTHAL, H.
Math. Zeitschr. Bd. 63, S. 374—405 (1956)

Neuere Fassungen des Riemann-Helmholtz-Lieschen Raumproblems*)

ISSAI SCHUR zum Gedächtnis

Von

HANS FREUDENTHAL

Einleitung

Riemanns Habilitationsvortrag von 1854 „Über die Hypothesen, welche der Geometrie zugrunde liegen" wurde erst 1867 von Dedekind veröffentlicht, und fast gleichzeitig erschien Helmholtz' Abhandlung „Über die Thatsachen, die der Geometrie zum Grunde liegen", die im Titel eine deutliche Spitze gegen Riemann verrät. Gegen Riemann, der als philosophischer Schüler Herbarts höchstens die *Topologie* des Raumes als unhypothetisch gelten läßt, und für den die *Metrik* schon Hypothese ist, wendet Helmholtz ein: Die bloße Tatsache einer Metrik erfordere schon einen Kongruenzbegriff für starre Körper. Denn wie könnte man — gäbe es keine starren Körper — Maße am selben Ort und an verschiedenen Orten vergleichen? Fundamental sei also nicht die von Riemann postulierte quadratische Maßbestimmung im Infinitesimalen, sondern die Kongruenz von starren Körpern, die dann bei der (Helmholtzschen) Lösung des Raumproblems die Existenz der quadratischen Maßbestimmung nach sich ziehe.

Natürlich ist Helmholtz' Einwand, den auch B. Erdmann mit anderen übernimmt, ungerechtfertigt. Denn die praktische Metrik erfordert nicht feste Körper, sondern nur feste Maßstäbe (also eindimensionale „Körper") als Meßgeräte, und mehr braucht Riemann nicht. Riemanns Ansatz, der von der quadratischen Differentialform ausgeht und hinterdrein, wenn es sein muß, den euklidischen oder nichteuklidischen Charakter der Geometrie durch Homogenitätsforderungen erzwingt, hat sich denn auch als fruchtbarer erwiesen als der Helmholtzsche, bei dem die Homogenität vorausgeschickt wird. Als Rechtfertigung der klassischen Metriken aber ist Helmholtz' Analyse effektvoller. Doch hat ihre erste wirklich befriedigende Formulierung fast hundert Jahre auf sich warten lassen.

Bei Helmholtz fehlt zunächst der Begriff der Gruppe, er wird immerzu implizit vorausgesetzt. Die Gruppe tritt bei ihm auf als Invarianzgruppe einer Funktion von zwei Punkten, die auf der gegebenen n-dimensionalen Mannigfaltigkeit variieren. Wir würden heutzutage dieser Funktion gleich die axiomatischen Forderungen auferlegen, der man eine Metrik unterwirft.

*) Die Resultate sind vorläufig mitgeteilt worden auf dem Internationalen Mathematikerkongreß, Amsterdam 1954, und auf der Berliner Riemann-Tagung 1954.

Hält man $k(\leq n-1)$ Punkte in allgemeiner Lage fest, so soll ein $(k+1)$-ter Punkt in allgemeiner Lage noch aller Ortsveränderungen fähig sein, die mit der Invariante verträglich sind; für $k=n$ ist keine stetige Ortsveränderung mehr zugelassen. Im Sinne dieser „freien Beweglichkeit der starren Körper" wird man Helmholtz' Hauptforderung (nach Lie) interpretieren, wenn man sich sehr große Freiheiten erlauben will. Aber ohne solche Freiheiten kann man mit ihnen überhaupt nichts anfangen, denn so exakt seine Formulierungen 1868 scheinen mochten, so verschwommen sahen sie zwanzig Jahre später aus, als Lie sie seiner Kritik unterwarf.

Helmholtz erschließt aus seinen Postulaten die euklidische und die nichteuklidischen Geometrien, die „klassischen" Geometrien des Raumproblems (anfangs übersieht er die hyperbolische — der Zusammenhang zwischen hyperbolischer Geometrie und den Räumen konstanter negativer Krümmung war im gleichen Jahre von Beltrami entdeckt worden).

In seiner Beweisführung nahm Helmholtz die Theorie der kontinuierlichen Gruppen vorweg. Seine Methode war aber ganz unzureichend, wie Lie bemerkte. Helmholtz verwendete nicht die „freie Beweglichkeit", die er postuliert, sondern die „freie Beweglichkeit im Infinitesimalen", die tiefer einschneidet. Diese „freie Beweglichkeit im Infinitesimalen" besagt: Man betrachte die Vereine, die bestehen aus einem Punkt, einer 1-Richtung, einer 2-Richtung, ..., einer $(n-1)$-Richtung (alle in Inzidenz); über der Menge dieser Vereine sei die Gruppe transitiv. — Diese Forderung führt recht einfach, wie Lie zeigt, zu den klassischen Gruppen des Raumproblems. (Siehe auch den Weylschen Beweis!)

Lie wertet auch das Postulat der freien Beweglichkeit der starren Körper aus. Genauer gesagt, begnügt er sich mit der Forderung, die Gruppe sei transitiv über der Menge der Vereine von $n-2$ Punkten in allgemeiner Lage (wenn n die Dimension der Mannigfaltigkeit ist). Er vermutet übrigens, daß man für $n>4$ mit weniger Punkten auskommt — eine Vermutung, die sich neuerdings bestätigt hat: für alle Dimensionen reicht „drei" aus.

Lie setzt natürlich voraus, daß er sich in einer genügend regulären Mannigfaltigkeit befindet, daß die Gruppe die nötigen Differenzierbarkeitseigenschaften besitzt; er bewegt sich immer im Kleinen und läßt sich durch topologische Schwierigkeiten, die man erst später erkannt hat, nicht beirren. Es ist schwer zu schätzen, wieviel von seinen Betrachtungen man heute noch gelten lassen will.

Der ständige Einwand, den Zeitgenossen gegen solche Betrachtungen erhoben, war der eines circulus vitiosus: in der „Zahlenmannigfaltigkeit" werde ja schon die ganze Geometrie postuliert (insbesondere auch die Unendlichkeit des Raumes, die hinterdrein in der sphärischen Geometrie wieder geleugnet werde). Auch noch 1872 — nachdem Cantor und Dedekind geometriefrei das Zahlenkontinuum definiert hatten und Weierstraß' Aufbau bekannt geworden war — hört man den Einwand noch sehr dringend, und erst seit Stolz' Wiederentdeckung der Bedeutung des Archimedischen Axioms (1881—1883) und Veroneses nicht-Archimedischer Geometrie (1890) nimmt

der Widerstand ab. Doch auch Hilberts „Grundlagen der Geometrie" begegnen ihm noch.

Im Anhang IV (1902) löst Hilbert das Raumproblem für die Ebene mit rein topologischen Mitteln. Als Substrat nimmt er die Zahlenebene (schließt also von vornherein elliptische und sphärische Geometrie aus). Von der Bewegungsgruppe F fordert er erstens, daß die Menge der Fx_1 mit $Fx_0 = x_0$ für $x_1 \neq x_0$ unendlich sei, ferner daß aus $f_n \in F$, $f_n(x_0) = x_n$, $f_n(y_0) = y_n$, $f_n(z_0) = z_n$ und $\lim x_n = x$, $\lim y_n = y'$, $\lim z_n = z'$ die Existenz eines $f' \in F$ folge mit $f(x_0) = x'$, $f(y_0) = y$, $f(z_0) = z'$. Durch diese zweite Forderung wird der Gruppe ungemerkt eine gewisse „Starrheit" auferlegt.

Partielle Behandlungen erfuhr das Raumproblem im weiteren Verlaufe von L. E. J. Brouwer (1909—1910), R. L. Moore (1919), Süß (1925—1927), St. S. Cairns (1923), R. G. Lubben (1928), Kérékjartó (1950), Montgomery und Zippin (1940).

Allgemein hat Kolmogorov es 1930 von neuem angegriffen. Er verlangt vom Raume R Metrisierbarkeit, Kompaktheit im Kleinen und Zusammenhang, von der Gruppe F Transitivität und eine Art von gleichmäßiger Stetigkeit, die man auch Starrheit nennen dürfte. Außerdem fordert er: Ist $x_0 \in R$, J seine Standuntergruppe ($f \in J$ bedeute: $f \in F$ und $fx_0 = x_0$), ist B eine Bahn ($= Jx_1$, $x_1 \neq x_0$) und B' eine zweite Bahn, so wird B durch B' oder B' durch B von x_0 getrennt. Auf den Bahnen sind wieder Bahnen zweiter Stufe definiert, die ebenfalls die Trennungseigenschaft besitzen sollen, usw. bis die Bahnen schließlich 0-dimensional werden.

Kolmogorov hat seinen Beweis für die Behauptung, daß die Forderungen zu den klassischen Lösungen des Raumproblems führen, nicht veröffentlicht; bis heute ist sie nur unter der zusätzlichen Voraussetzung der Vollständigkeit von F bewiesen.

Inzwischen hatten Garret Birkhoff (1941, 1944) und H. Busemann (1941, 1942) die „Invariante" der klassischen Untersuchungen wieder in den Gesichtskreis bezogen. Sie betrachteten metrische Räume und stellten Forderungen an die Gruppe der Isometrien: Jede Isometrie von Teilmengen (Birkhoff) bzw. von dreipunktigen Teilmengen (Busemann) soll zu einer Isometrie des ganzen Raumes fortgesetzt werden können. Außerdem stellten sie noch Konvexheits- und ähnliche Forderungen. Einen weiteren Schritt tat H. C. Wang (1951, 1952), der die kompakten, zusammenhängenden metrischen Räume klassifiziert, in denen jedes Punktepaar in jedes kongruente durch eine Isometrie des Raumes übergeführt werden kann. Auch die im kleinen kompakten Räume behandelte Wang, jedoch unter zusätzlichen Voraussetzungen (darunter die einer geraden Dimension). Die inzwischen erzielte Lösung von Hilberts 5. Problem greift in Wangs Untersuchungen entscheidend ein.

In einer kurzen Note (ohne Beweise) hat J. Tits (1952) die Untersuchungen zum Raumproblem zu einem gewissen Abschluß gebracht. Er ist wieder zu Kolmogorovs Ansatz zurückgekehrt. Er setzt aber von Kolmogorovs Trennungseigenschaften nur die erste voraus und klassifiziert die Räume und

Gruppen, die die Forderungen befriedigen[1]). In seinen Ergebnissen ist das von Wang enthalten. Die Trennungseigenschaft der nächsten Stufe führt bereits zu den klassischen Räumen.

Ich bin zu der vorliegenden Arbeit durch die Titssche Note angeregt worden. Ich komme mit schwächeren Voraussetzungen als Tits aus. Meine Beweismethode ist ganz verschieden von der von Tits[2]) und, wie mir scheint, einfacher.

Ich werde nun die Titssche Fassung des Raumproblems mit meiner Verschärfung auseinandersetzen. Wir verlangen vom Raum R

Kompaktheit im Kleinen und Zusammenhang.

In R wirkt eine transitive Gruppe von Homöomorphismen, die die Eigenschaften S, V, Z besitzt:

S (Starrheit): Zu je zwei abgeschlossenen Mengen A, B von R mit $A \cap B = O$, von denen mindestens eine kompakt ist, gibt es eine offene Menge $U(\neq O)$, so daß für jedes $f \in F$ die Beziehungen $f(U) \cap A \neq O$ und $f(U) \cap B \neq O$ einander ausschließen.

Das ist im wesentlichen die Forderung, der Kolmogorov den Namen „gleichmäßige Stetigkeit" gibt. Man könnte sie übrigens noch abschwächen. Sie führt zu einer gleichmäßigen Struktur in R, die bei F invariant ist. Im Sinne dieser Struktur fordert man nun

V: Vollständigkeit von F.

Es wäre interessant festzustellen, in welchem Maße diese Forderung im weiteren Verlaufe wirklich nötig ist.

Ist $x_0 \in R$, so heißt *Standuntergruppe* J von x_0 die Gruppe der $f \in F$ mit $f x_0 = x_0$. Jede Menge $J x_1$ heißt *Bahn*.

Z: Es gibt eine Bahn, die R zerlegt.

Hier stellt Tits (wie Kolmogorov) die schärfere Forderung, daß unter je zwei Bahnen $(\neq (x_0))$ eine ist, die zwischen x_0 und der anderen den Raum zerlegt.

Durch die Bedingungen S, V, Z werden die Möglichkeiten sehr stark beschränkt. Mit allgemein gruppentheoretisch-topologischen Methoden (unter anderem unter Verwendung der Yamabeschen Lösung des verallgemeinerten fünften Hilbertschen Problems[3])) gelangt man zu Satz I, der aussagt, daß F notwendig eine (eventuell gemischte) transitive Lie-Gruppe mit endlich vielen Komponenten ist, und daß J kompakt und (bei geeigneter Koordinatenwahl) eine lineare orthogonale und auf der Abstandssphäre transitive Gruppe ist.

Die Darstellungstheorie der halbeinfachen Gruppen liefert dann die Methoden zur Klassifikation der zulässigen Standuntergruppen J und zum Einbau dieser in die möglichen F. Das endgültige Ergebnis können wir zusammenfassen in einer (von Tits herrührenden) Liste.

[1]) Die Klassifikation ist unvollständig, aber später korrigiert worden.

[2]) die ich inzwischen aus einem Manuskript kennengelernt habe.

[3]) Yamabe, H.: On the conjecture of Iwasawa and Gleason. Ann. of Math. (2) **58**, 48—54 (1953). — Yamabe, H.: A generalization of a theorem of Gleason. Ann. of Math. (2) **58**, 351—365 (1953).

Ehe wir diese angeben, führen wir einige Bezeichnungen ein.

Mit $*$ geben wir einen der folgenden Körper an:

re: Körper der reellen Zahlen.

ko: Körper der komplexen Zahlen.

qu: (Schief-)Körper der Quaternionen (über re).

ok: (Alternativ-)Körper der Oktaven (über re).

$\mathrm{VR}(n, *)$ ist der n-dimensionale Vektorraum über $*$. (Für $* = \mathrm{ok}$ wird nur $n = 2$ verwendet.)

$\mathrm{AR}(n, *)$ ist der entsprechende affine Raum [d.h. $\mathrm{VR}(n, *)$ nach Aufhebung der Auszeichnung von 0].

$\mathrm{UVR}(n, *)$ und $\mathrm{UAR}(n, *)$ sind dieselben Räume nach Adjunktion einer hermiteschen Form $\sum x_i \bar{x}_i$ in den Koordinaten (unitäre Räume).

$\mathrm{PR}(n, *)$ ist der n-dimensionale projektive Raum, der aus $\mathrm{VR}(n+1, *)$ entsteht als Raum der Klassen (ϱx) $(\varrho \in *,\ x \in \mathrm{VR})$. Im Falle ok muß PR anders definiert werden[4]).

$\mathrm{ElR}(n, *)$ ist der n-dimensionale elliptische Raum und entsteht aus $\mathrm{PR}(n, *)$ durch Adjunktion einer Absoluten $\sum x_i \bar{x}_i = 0$. Im Falle ok muß ElR anders definiert werden[4]).

$\mathrm{SpR}(n, \mathrm{re})$ ist der n-dimensionale sphärische Raum, der als Überlagerung von $\mathrm{ElR}(n, \mathrm{re})$ entsteht, oder als $\sum_1^{n+1} x_i^2 = 1$ in $\mathrm{VR}(n+1, \mathrm{re})$.

$\mathrm{HyR}(n, *)$ ist das Innere einer Absoluten $\sum_1^{n} x_i \bar{x}_i = x_{n+1} \bar{x}_{n+1}$ in $\mathrm{PR}(n, *)$. Im Falle ok muß HyR anders definiert werden.

Zu diesen Räumen gehören die Gruppen der (stetigen) Automorphismen, die durch ein vorgesetztes G bezeichnet werden. Insbesondere sind

GUAR	die euklidischen	Bewegungsgruppen.
GElR	die elliptischen	
GSpR	die sphärischen	
GHyR	die hyperbolischen	

Dabei ist $\mathrm{GSpR}(n, \mathrm{re})$ die Gruppe $\mathrm{GUVR}(n+1, \mathrm{re})$ auf $\sum_1^{n+1} x_i^2 = 1$ betrachtet.

Von den Gruppen GUAR betrachten wir noch gewisse Untergruppen G_0, die definiert sind

für $* = \mathrm{re}, \mathrm{ko}$ durch: Funktionaldeterminante $= 1$,

für $* = \mathrm{qu}$, durch $b = 1$ in der Formel

$$x_i' = \sum_{j=1}^{n} a_{ij} x_j b + c_i$$

für ein Element von $\mathrm{GUAR}(n, \mathrm{qu})$.

[4]) Siehe FREUDENTHAL, H.: Oktaven, Ausnahmegruppen und Oktavengeometrie. Utrecht 1951.

Ferner verstehe man für $\mathrm{ElR}(n, \mathrm{re})$, $\mathrm{ElR}(n, \mathrm{ko})$ (n ungerade), $\mathrm{SpR}(n, \mathrm{re})$ ($n > 1$) unter G_0 die Komponente der Identität von G.

Übrigens sei $G_0 = G$.

Hauptsatz I. *Die den Bedingungen S, V, Z genügenden Gruppen F sind: Die zwischen* G *und* G_0 *von*

$$\mathrm{UAR}(n, *), \quad \mathrm{ElR}(n, *), \quad \mathrm{Sp}(n, \mathrm{re}), \quad \mathrm{Hy}(n, *)$$

liegenden abgeschlossenen Gruppen.

Ferner (1) *die Untergruppe von* $\mathrm{GUAR}(7, \mathrm{re})$, *deren Standuntergruppe eine Darstellung der* G_2 *(bzw. ergänzt mit einer Spiegelung), nämlich die Automorphismengruppe der Oktaven mit Realteil* 0, *ist;*

(2) *die Untergruppe von* $\mathrm{GUAR}(8, \mathrm{re})$, *deren Standuntergruppe die niedrigste Spindarstellung von* $\mathrm{GUVR}(7, \mathrm{re})$ *ist;*

(3) *die Untergruppe von* $\mathrm{GSpR}(6, \mathrm{re})$, *die entsteht, wenn man obengenannte Darstellung der* G_2 *auf* $\mathrm{SpR}(6, \mathrm{re})$ *betrachtet (eventuell ergänzt mit einer Spiegelung);*

(4) *die Untergruppe von* $\mathrm{GSpR}(7, \mathrm{re})$, *die entsteht, wenn man oben genannte Spindarstellung von* $\mathrm{GUVR}(7, \mathrm{re})$ *auf* $\mathrm{SpR}(7, \mathrm{re})$ *betrachtet;*

(5, 6) *die den beiden letzten Gruppen* (3—4) *entsprechenden Gruppen von* $\mathrm{PR}(6, \mathrm{re})$ *und* $\mathrm{PR}(7, \mathrm{re})$.

Man sieht also, daß schon Kolmogorovs Trennungseigenschaft erster Stufe tief einschneidet, und daß man sie nur für eine einzige Bahn zu fordern braucht. Will man nun die klassischen Räume des Riemann-Helmholtz-Lieschen Raumproblems charakterisieren, so wird man eine Trennungseigenschaft höherer Stufe fordern, und das bedeutet, daß eine Bahn (mit der auf ihr herrschenden Darstellung der Standuntergruppe) wieder ein S, V, Z genügender Raum sein muß, und so fort.

Aus der Liste von Hauptsatz I folgt:

Hauptsatz II. *Fordert man die Gültigkeit von Z auch für eine Bahn, so bleiben übrig die klassischen Gruppen des Raumproblems*

$$\mathrm{GUAR}(n, \mathrm{re}), \quad \mathrm{GElR}(n, \mathrm{re}), \quad \mathrm{GSpR}(n, \mathrm{re}), \quad \mathrm{GHyR}(n, \mathrm{re})$$

und die zugehörigen Gruppen G_0; *ferner* (1), (2), (4), (6), *auf deren Bahnen bzw. die Gruppen* (3), (4), (3), (3) *herrschen. Die zugehörigen Räume sind alle von konstanter Krümmung. Fordert man Z auch noch für die Bahnen zweiter Stufe, so bleibt außer den klassischen Gruppen nur* (2) *übrig. Fordert man schließlich Z noch für die Bahnen dritter Stufe, so bleiben nur die klassischen Gruppen übrig.*

Verfolgt man die Dinge weiter in der Richtung von G. Birkhoff, Busemann und Wang, so erhält man:

Hauptsatz III. *R sei metrisch, zusammenhängend und im Kleinen kompakt. F sei die Gruppe der Isometrien von R. Es gebe eine Zahl* $\gamma > 0$, *aber*

kleiner als die obere Grenze der in R auftretenden Abstände, mit der Eigenschaft: F führt ein Punktepaar mit dem Abstand γ in jedes andere mit dem Abstand γ über. Dann ist F eine der Gruppen

$$\mathrm{GUAR}(n, \mathrm{re}),\quad \mathrm{GElR}(n, *),\quad \mathrm{GSpR}(n, *),\quad \mathrm{GHyR}(n, *).$$

Hauptsatz IV. *Es gebe außerdem eine Zahl $\gamma' > 0$, aber kleiner als der Durchmesser der Abstandssphäre mit dem Radius γ, mit der Eigenschaft, daß F ein Tripel mit den Abständen γ, γ, γ' in jedes andere derartige überführe. Dann ist F eine der Gruppen*

$$\mathrm{GUAR}(n, \mathrm{re}),\quad \mathrm{GElR}(n, \mathrm{re}),\quad \mathrm{GSpR}(n, \mathrm{re}),\quad \mathrm{GHyR}(n, \mathrm{re}).$$

Literatur zur Einleitung

RIEMANN, B.: Über die Hypothesen, welche der Geometrie zugrunde liegen, 1854 gehaltener Habilitationsvortrag, 1867 veröffentlicht in Abh. Ges. Wiss. Göttingen **15** = Werke 1876, 254—269.

HELMHOLTZ, H.: Über die thatsächlichen Grundlagen der Geometrie. (Vortrag vom 22. Mai 1866?) Verh. Naturhist. med. Verein Heidelberg **4**, 197—202 (1868) = Wissensch. Abh. II (1883), 610—617.

HELMHOLTZ, H.: Über die Thatsachen, die der Geometrie zum Grunde liegen. Nachr. Ges. Wiss. Göttingen **1868**, 193—221 = Wissensch. Abh. II (1883), 618—639.

ERDMANN, B.: Die Axiome der Geometrie. Leipzig 1877.

LIE, S.: Über die Grundlagen der Geometrie. Berichte Ges. Wiss. Leipzig **42**, 284—321, 355—418 (1890) = Ges. Abh. II, 1, 380—468.

LIE, S.: Theorie der Transformationsgruppen III, Abt. V. Leipzig 1893.

CANTOR, G.: Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math. Ann. **5**, 123—132 (1872) = Ges. Abh. 92—102.

DEDEKIND, R.: Stetigkeit und irrationale Zahlen. Braunschweig 1872.

STOLZ, O.: B. Bolzano's Bedeutung in der Geschichte der Infinitesimalrechnung. Math. Ann. **18**, 255—279 (1881), speziell 269.

STOLZ, O.: Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes. Math. Ann. **22**, 504—519 (1883).

VERONESE, G.: I continuo rettilineo e l'assioma V d'Archimede. Atti Accad. Linc., Memorie cl. fis. (4) **6**, 603—624 (1890).

HILBERT, D.: Grundlagen der Geometrie, Anhang IV (1902).

WEYL, H.: Mathematische Analyse des Raumproblems. Berlin 1923.

BROUWER, L. E. J.: Die Theorie der endlichen kontinuierlichen Gruppen, unabhängig von den Axiomen von Lie. Math. Ann. **67**, 246—267 (1909); **69**, 181—203 (1910).

MOORE, R. L.: On the Lie-Riemann-Helmholtz-Hilbert problem of the foundations of geometry. Amer. J. Math. (2) **41**, 299—319 (1919).

SÜSS, W.: Beiträge zur gruppentheoretischen Begründung der Geometrie. Tohoku Math. J. **26**, 365—385 (1926); **27**, 214—242 (1926); **28**, 228—241 (1927). — Jap. J. Math. **4**, 97—119 (1927).

CAIRNS, ST. S.: An axiomatic basis for plane geometry. Trans. Amer. Math. Soc. **35**, 234—244 (1923).

LUBBEN, R. G.: The double elliptic case of the Lie-Riemann-Helmholtz problem of the foundation of geometry. Fundamenta Math. **11**, 34—95 (1928).

KÉRÉKJARTO, B. DE: Sur le groupe des homographies et des antihomographies d'une variable complexe. Comment. Helvet. **13**, 68—82 (1950).

MONTGOMERY, D., and L. ZIPPIN: Topological group foundations of rigid space geometry. Trans. Amer. Math. Soc. **48**, 21—49 (1940).

KOLMOGOROFF, A.: Zur topologisch-gruppentheoretischen Begründung der Geometrie. Nachr. Ges. Wiss. Göttingen **1930**, 208—210.

BIRKHOFF, GARRET: Metric foundations of Geometry. Proc. Nat. Acad. Sci. **27**, 402—406 (1941).

BIRKHOFF, GARRET: Metric foundations of Geometry I. Trans. Amer. Math. Soc. **55**, 465—492 (1944).

BUSEMANN, H.: Metric Methods in Finsler spaces. Ann. of Math. Studies **8** (1942).

WANG, H. C.: Two-point homogeneous spaces. Ann. of Math. (2) **55**, 177—191 (1952).

WANG, H. C.: Two theorems on metric spaces. Pacific J. Math. **1**, 473—480 (1951).

TITS, J.: Etude de certains espaces métriques. Bull. Soc. Math. Belg. **1952**, 44—52 (1953).

TITS, J.: Sur un article précédent: Etudes de certains espaces métriques. Bull. Soc. Math. Belg. **1953**, 124—125 (1953).

Vorbemerkungen

1. „kompakt" wird im Sinne des früheren „bikompakt" gebraucht.

2. Wo von Gruppen die Rede ist, handelt sich es meistens um Transformationsgruppen, d.h. um in gewisser Weise dargestellte Gruppen. Man soll also z.B. die Bezeichnung GUVR$(n$, re$)$ nicht verwenden wie $O(n)$, das immer mehr benutzt wird, um die n-dimensionale Drehgruppe zu bezeichnen und nicht ihre Darstellung im n-Raum, für die wir gerade den Ausdruck GUVR$(\mathfrak{g}$, re$)$ eingeführt haben, um Mißverständnissen vorzubeugen.

3. Der Infinitesimalring einer Lie-Gruppe wird durch den entsprechenden fetten Buchstaben bezeichnet; z.B. Infinitesimalring von F ist $\boldsymbol{F}$, von B_3 ist $\boldsymbol{B}_3$. Auch Teilmengen und Elemente von Infinitesimalringen werden fett bezeichnet.

4. Die einfachen Lie-Gruppen (Lie-Algebren) werden gemäß der Cartanschen Klassifikation bezeichnet. Hinzugefügt ist: $\boldsymbol{A}_0$ für die 1-dimensionale Algebra, $\boldsymbol{O}$ für die 0-dimensionale. Auch für die fundamentalen Darstellungen bedienen wir uns der Cartanschen[5]) Bezeichnungen.

5. *Gemischte* Lie-Gruppe ist die Bezeichnung für eine Gruppe mit einer diskreten Menge von Komponenten und mit einer *reinen* (d.h. zusammenhängenden) Lie-Gruppe als Komponente der Eins.

6. In einer Transformationsgruppe bedeute id die identische Transformation.

1. Topologische Starrheit

1.1. *Räume* werden im folgenden *Hausdorffsch, im Kleinen kompakt* und *zusammenhängend* vorausgesetzt.

Im Raum R sei eine *transitive Gruppe* F von *Homöomorphismen* (von R auf sich) gegeben. Der Gruppe F erlegen wir noch die Bedingung auf:

S (topologische Starrheit): Zu je zwei fremden abgeschlossenen Mengen A und B von R, von denen mindestens eine kompakt ist, gibt es eine nichtleere offene Menge U, so daß für jedes $f \in F$ die Beziehungen $fU \cap A \neq O$, $fU \cap B \neq O$ einander ausschließen.

[5]) Oeuvres I, 1, 355—398.

1.2. Auf Grund von (S) versieht man R mit einer invarianten gleichmäßigen Struktur:

Ist U offen in R, so sei $\tilde{U}$ die Menge der $(x, y) \in R \times R$ mit der Eigenschaft: es gibt ein $f \in F$ mit $x, y \in fU$. Diese $\tilde{U}$ spielen die Rolle eines Umgebungssystems der Diagonale von $R \times R$.

Unter $\tilde{U}_1 \tilde{U}_2$ versteht man bekanntlich die Menge der (x, z) mit der Eigenschaft: es gibt ein y mit $(x, y) \in \tilde{U}_1$, $(y, z) \in \tilde{U}_2$. Wir zeigen:

Zu jedem V (offen $\neq O$) gibt es ein U (offen $\neq O$) mit $\tilde{U}^2 \subset \tilde{V}$.

Beweis. Zu V gibt es ein nichtleeres offenes kompaktes V_1 mit $\overline{V}_1 \subset V$. Auf $\overline{V}_1$ und R/V wenden wir (S) an: es gibt ein U_1, so daß aus $f \in F$, $fU_1 \cap \overline{V}_1 \neq O$ folgt: $fU_1 \subset V$. Wegen der Transitivität von F dürfen wir annehmen: $U_1 \cap V_1 \neq O$. Wir setzen $U_1 \cap V_1 = U$. Dann gilt auch:

$$\text{aus} \quad fU \cap U \neq O \quad \text{folgt} \quad fU \subset V.$$

Sei nun (x, y), $(y, z) \in \tilde{U}$, also mit geeigneten $f_i \in F$:

$$x, y \in f_1 U, \quad y, z \in f_2 U,$$

also
$$f_1 U \cap f_2 U \neq O,$$
$$f_2^{-1} f_1 U \cap U \neq O,$$
$$f_2^{-1} f_1 U \subset V,$$
$$f_1 U \subset f_2 V,$$

aber auch
$$f_2 U \subset f_2 V.$$

Also
$$x, z \in f_2 V,$$
$$(x, z) \in \tilde{V}.$$

Also $\tilde{U}^2 \subset \tilde{V}$, w. z. b. w.

1.3. Es gibt ein nichtleeres offenes $U \subset R$ mit der Eigenschaft: ist $A \subset R$ kompakt, so ist auch

$$\cup (fU \mid f \in F, fU \cap A \neq O) \quad \text{kompakt.}$$

Beweis. Wir wählen V und U, so daß V kompakt und $\tilde{U}^2 \subset \tilde{V}$ ist. Wir überdecken $\bar{A}$ mit endlich vielen $f_i U$ $(f_i \in F)$. Zu jedem $f \in F$ mit $fU \cap A \neq O$ gibt es ein i mit $fU \cap f_i U \neq O$. Also $f_i^{-1} fU \cap U \neq O$, $f_i^{-1} fU \subset V$, $fU \subset f_i V$. Die Vereinigung dieser fU ist also in der der $f_i V$, also kompakt.

1.4. U sei gemäß 1.3 gewählt. Sei W_n die $\tilde{U}^n$-Umgebung irgendeines $x_0 \in R$. Dann ergibt sich induktiv: W_n ist kompakt. W_n ist nämlich die Menge der z mit $y \in W_{n-1}$, $(y, z) \in \tilde{U}$, also die Vereinigung der fU mit $fU \cap W_{n-1} \neq O$. Nach 1.3 folgt aus der Kompaktheit von W_{n-1} die von W_n.

1.5. Weiter gilt: $\cup W_n = R$.

Beweis. R ist zusammenhängend, $\cup W_n$ ist offen. Wir zeigen, daß es auch abgeschlossen ist: wäre $x \in \overline{\cup W_n} \setminus \cup W_n$, so gälte für jedes f mit $x \in fU : fU \cap \cup W_n \neq \dot{O}$. Dann gäbe es ein i mit $fU \cap W_i \neq O$, also $fU < W_{i+1}$, $x \in W_{i+1}$ im Widerspruch zur Annahme.

1.6. Wir definieren eine Topologie in F:

Für $U \neq O$ offen, A kompakt in R und $f_i \in F$ setzen wir

$$(f_1, f_2) \in \{U, A\}$$

dann und nur dann, wenn

$$(f(x), g(x)) \in \tilde{U} \text{ für alle } x \in A.$$

Die $\{U, A\}$ gelten als Umgebungssystem der Diagonale von $F \times F$. Unter $\{U, A\}$-Umgebung von f_1 verstehen wir die Menge der f_2 mit $(f_1, f_2) \in \{U, A\}$.

Das ist eine Hausdorffsche Topologie[6]).

1.7. F wird mit dieser Topologie eine topologische Gruppe:

Zu gegebenen $f_i \in F$ mit

$$f_1 f_2 = f_3$$

und gegebenem $\{V, A\}$ bestimmen wir $\{U, A_1\}$ und $\{U, A_2\}$, so daß aus $(f_i, g_i) \in \{U, A_i\}$ $(i = 1, 2)$ folgt $(f_3, g_3) \in \{V, A\}$. Wir wählen nämlich U, so daß

$$\tilde{U}^2 < \tilde{V}$$

und

$$A_1 = f_2 A, \quad A_2 = A.$$

Dann wird

$$(f_1 x, g_1 x) \in \tilde{U} \quad \text{für} \quad x \in f_2 A,$$

$$(f_1 f_2 x, g_1 f_2 x) \in \tilde{U} \quad \text{für} \quad x \in A.$$

Weiter

$$(f_2 x, g_2 x) \in \tilde{U} \quad \text{für} \quad x \in A,$$

$$(g_1 f_2 x, g_1 g_2 x) \in \tilde{U} \quad \text{für} \quad x \in A.$$

Also

$$(f_1 f_2 x, g_1 g_2 x) \in \tilde{U}^2 < V \quad \text{für} \quad x \in A, \quad \text{w. z. b. w.}$$

Ferner ergibt sich mit

$$(f, g) \in \{V, f^{-1} A\}:$$

$$(f x, g x) \in \tilde{V} \quad \text{für} \quad x \in f^{-1} A,$$

$$(x, g f^{-1} x) \in \tilde{V} \quad \text{für} \quad x \in A,$$

$$(g^{-1} x, f^{-1} x) \in \tilde{V} \quad \text{für} \quad x \in A,$$

und damit ist auch die Stetigkeit von $f \to f^{-1}$ erwiesen.

[6]) Siehe BOURBAKI, l. III, chap. X, § 1, no. 5, prop. 2.

1.2. Auf Grund von (S) versieht man R mit einer invarianten gleichmäßigen Struktur:

Ist U offen in R, so sei $\tilde{U}$ die Menge der $(x, y) \in R \times R$ mit der Eigenschaft: es gibt ein $f \in F$ mit $x, y \in fU$. Diese $\tilde{U}$ spielen die Rolle eines Umgebungssystems der Diagonale von $R \times R$.

Unter $\tilde{U}_1 \tilde{U}_2$ versteht man bekanntlich die Menge der (x, z) mit der Eigenschaft: es gibt ein y mit $(x, y) \in \tilde{U}_1$, $(y, z) \in \tilde{U}_2$. Wir zeigen:

Zu jedem V (offen $\neq O$) gibt es ein U (offen $\neq O$) mit $\tilde{U}^2 \subset \tilde{V}$.

Beweis. Zu V gibt es ein nichtleeres offenes kompaktes V_1 mit $\overline{V}_1 \subset V$. Auf $\overline{V}_1$ und R/V wenden wir (S) an: es gibt ein U_1, so daß aus $f \in F$, $fU_1 \cap \overline{V}_1 \neq O$ folgt: $fU_1 \subset V$. Wegen der Transitivität von F dürfen wir annehmen: $U_1 \cap V_1 \neq O$. Wir setzen $U_1 \cap V_1 = U$. Dann gilt auch:

$$\text{aus} \quad fU \cap U \neq O \quad \text{folgt} \quad fU \subset V.$$

Sei nun (x, y), $(y, z) \in \tilde{U}$, also mit geeigneten $f_i \in F$:

$$x, v \in f_1 U, \quad y, z \in f_2 U,$$

also
$$f_1 U \cap f_2 U \neq O,$$
$$f_2^{-1} f_1 U \cap U \neq O,$$
$$f_2^{-1} f_1 U \subset V,$$
$$f_1 U \subset f_2 V,$$

aber auch
$$f_2 U \subset f_2 V.$$

Also
$$x, z \in f_2 V,$$
$$(x, z) \in \tilde{V}.$$

Also $\tilde{U}^2 \subset \tilde{V}$, w. z. b. w.

1.3. Es gibt ein nichtleeres offenes $U \subset R$ mit der Eigenschaft: ist $A \subset R$ kompakt, so ist auch

$$\cup (fU \mid f \in F,\ fU \cap A \neq O) \quad \text{kompakt.}$$

Beweis. Wir wählen V und U, so daß V kompakt und $\tilde{U}^2 \subset \tilde{V}$ ist. Wir überdecken $\bar{A}$ mit endlich vielen $f_i U$ $(f_i \in F)$. Zu jedem $f \in F$ mit $fU \cap A \neq O$ gibt es ein i mit $fU \cap f_i U \neq O$. Also $f_i^{-1} fU \cap U \neq O$, $f_i^{-1} fU \subset V$, $fU \subset f_i V$. Die Vereinigung dieser fU ist also in der der $f_i V$, also kompakt.

1.4. U sei gemäß 1.3 gewählt. Sei W_n die $\tilde{U}^n$-Umgebung irgendeines $x_0 \in R$. Dann ergibt sich induktiv: W_n ist kompakt. W_n ist nämlich die Menge der z mit $y \in W_{n-1}$, $(y, z) \in \tilde{U}$, also die Vereinigung der fU mit $fU \cap W_{n-1} \neq O$. Nach 1.3 folgt aus der Kompaktheit von W_{n-1} die von W_n.

1.5. Weiter gilt: $\cup W_n = R$.

Beweis. R ist zusammenhängend, $\cup W_n$ ist offen. Wir zeigen, daß es auch abgeschlossen ist: wäre $x \in \overline{\cup W_n} \setminus \cup W_n$, so gälte für jedes f mit $x \in fU : fU \cap \cup W_n \neq 0$. Dann gäbe es ein i mit $fU \cap W_i \neq 0$, also $fU < W_{i+1}$, $x \in W_{i+1}$ im Widerspruch zur Annahme.

1.6. Wir definieren eine Topologie in F:

Für $U \neq 0$ offen, A kompakt in R und $f_i \in F$ setzen wir

$$(f_1, f_2) \in \{U, A\}$$

dann und nur dann, wenn

$$(f(x), g(x)) \in \tilde{U} \text{ für alle } x \in A.$$

Die $\{U, A\}$ gelten als Umgebungssystem der Diagonale von $F \times F$. Unter $\{U, A\}$-Umgebung von f_1 verstehen wir die Menge der f_2 mit $(f_1, f_2) \in \{U, A\}$.

Das ist eine Hausdorffsche Topologie[6]).

1.7. F wird mit dieser Topologie eine topologische Gruppe:

Zu gegebenen $f_i \in F$ mit

$$f_1 f_2 = f_3$$

und gegebenem $\{V, A\}$ bestimmen wir $\{U, A_1\}$ und $\{U, A_2\}$, so daß aus $(f_i, g_i) \in \{U, A_i\}$ $(i = 1, 2)$ folgt $(f_3, g_3) \in \{V, A\}$. Wir wählen nämlich U, so daß

$$\tilde{U}^2 < \tilde{V}$$

und

$$A_1 = f_2 A, \quad A_2 = A.$$

Dann wird

$$(f_1 x, g_1 x) \in \tilde{U} \quad \text{für} \quad x \in f_2 A,$$

$$(f_1 f_2 x, g_1 f_2 x) \in \tilde{U} \quad \text{für} \quad x \in A.$$

Weiter

$$(f_2 x, g_2 x) \in \tilde{U} \quad \text{für} \quad x \in A,$$

$$(g_1 f_2 x, g_1 g_2 x) \in \tilde{U} \quad \text{für} \quad x \in A.$$

Also

$$(f_1 f_2 x, g_1 g_2 x) \in \tilde{U}^2 < V \quad \text{für} \quad x \in A, \quad \text{w. z. b. w.}$$

Ferner ergibt sich mit

$$(f, g) \in \{V, f^{-1} A\}:$$

$$(f x, g x) \in \tilde{V} \quad \text{für} \quad x \in f^{-1} A,$$

$$(x, g f^{-1} x) \in \tilde{V} \quad \text{für} \quad x \in A,$$

$$(g^{-1} x, f^{-1} x) \in \tilde{V} \quad \text{für} \quad x \in A,$$

und damit ist auch die Stetigkeit von $f \to f^{-1}$ erwiesen.

[6]) Siehe Bourbaki, l. III, chap. X, § 1, no. 5, prop. 2.

1.8. Sei $x_0 \in R$. Dann gibt es eine Umgebung U von x_0, so daß die $\{U, x_0\}$-Umgebung der Identität von F kompakt ist im Raum der stetigen Abbildungen von R in R (im Sinne der kompakten Konvergenz)[7]).

Das folgt aus dem Satz von Ascoli[8]). Um ihn anwenden zu können, müssen wir noch zeigen, daß für jedes $x \in R$ die Menge der fx mit $(f, \mathrm{id}) \in \{U, x_0\}$ kompakt ist; U wählen wir dabei gemäß 1.3—5. Die Bedingung für f lautet dann auch

$$(fx_0, x_0) \in \tilde{U}.$$

Nach 1.5 gibt es ein i mit

$$(x_0, x) \in \tilde{U}^i.$$

$$(fx_0, fx) \in \tilde{U}^i,$$

$$(x_0, fx) \in \tilde{U}^{i+1},$$

und das ist nach 1.4 kompakt.

1.9. Wir setzen im folgenden noch voraus:

V: F ist vollständig.

Dann folgt aus 1.8: F ist im Kleinen kompakt.

1.10. Im Beweise von 1.8 können wir statt U irgendeine der Mengen W_n wählen (s. 1.4). Dann ergibt sich, daß die Menge der f mit

$$(f, \mathrm{id}) \in \{W_n, x_0\}$$

kompakt ist. Nach 1.5 ist F die Vereinigung der abzählbar vielen (kompakten) $\{W_n, x_0\}$-Umgebungen der Identität. Hieraus folgt:

Ist $Q \neq O$ offen in F, so ist F Vereinigung abzählbar vieler fQ $(f \in F)$.

1.11. Die kanonische Abbildung $\varphi: f \to fx_0$ von F auf R ist offen.

Dieser Satz wird im wesentlichen genau so bewiesen wie entsprechende Sätze, bei denen das zweite Abzählbarkeitsaxiom vorausgesetzt wird[9]). 1.10 kann nämlich dieses Axiom vertreten. Wir skizzieren den Beweis:

Sei $Q \neq O$ kompakt und offen in F. Man kann F mit abzählbar vielen fQ überdecken. Sei Qx_0 nirgends dicht. Dann ist auch fQx_0 nirgends dicht. R wird dann mit abzählbar vielen nirgends dichten $f_n Q x_0$ überdeckt. Sei $U_1 \neq O$ offen und kompakt in R. Wir definieren induktiv ein nichtleeres offenes $U_{n+1} < U_n$, so daß $\overline{U}_{n+1} \cap f_n Q x_0 = O$ ist; wegen der Nirgendsdichtheit ist das möglich. Wegen der Kompaktheit von U_1 ist $\cap \overline{U}_n$ nichtleer und weiter fremd zu allen $f_n Q x_0$, also zu R, und das ist ein Widerspruch.

Also ist Qx_0 irgendwo dicht. $\overline{Q}x_0 = \overline{Qx_0}$ enthält also einen inneren Punkt. Ist $Q_1 \neq O$ offen und $\overline{Q}_1 < Q$, so gilt analog: $\overline{Q}_1 x_0$ enthält einen inneren Punkt. Also auch: Qx_0 enthält einen inneren Punkt, also eine offene Menge V. Sei P die Menge der $f \in F$ mit $fx_0 \in V$. Dann ist P offen in F, also auch $P \cap Q = Q'$ offen in F und $Q'x_0 = V$ offen.

[7]) BOURBAKI, l. III, chap. X, § 1, no. 3.

[8]) BOURBAKI, l. III, chap. X, § 4, no. 1.

[9]) FREUDENTHAL, H.: Ann. of Math. **37**, Nr. 4—5, 46—56 (1936).

Also gibt es zu jedem offenen Q in F ein nichtleeres offenes $Q' \subset Q$, so daß $Q' x_0$ auch offen ist.

Sei P nun beliebig offen in F. Wir müssen zeigen: $P x_0$ offen in R. Wir dürfen annehmen: $\mathrm{id} \in P$. Es genügt auch, zu zeigen, daß x_0 innerer Punkt von $P x_0$ ist. Wir wählen das offene Q so, daß $Q^{-1} Q \subset P$ ist. Sei $f \in Q'$ (siehe vorigen Absatz). $f x_0$ ist dann innerer Punkt von $Q' x_0$, also x_0 innerer Punkt von $f^{-1} Q' x_0 \subset P x_0$, w. z. b. w.

2. Bahnzerlegung

2.1. Zu einem Punkt $x_0 \in R$ gehört die *Standuntergruppe* J von F, die besteht aus den $f \in F$ mit $f x_0 = x_0$. Nach 1.11 darf R mit dem Nebengruppenraum F/J kanonisch identifiziert werden. Jede Menge Jx (mit festem x) heißt eine *J-Bahn* oder kurz eine *Bahn*.

Ist U eine Umgebung von x_0 gemäß 1.3, so ist

$$\cup (fU \mid f \in J)$$

eine bei J invariante Umgebung von x_0. Sie ist in der $\tilde{U}^2$-Umgebung von x_0 enthalten, also nach 1.4 kompakt.

Ebenso ist nach 1.4—5 jede Bahn kompakt.

2.2. Wir stellen nun noch die Forderung:

Z: Es gibt eine Bahn, die R zerlegt.

Wir beweisen in 2.3—7:

Satz I. *Unter den Voraussetzungen (S, V, Z) ist F eine (eventuell gemischte) transitive Lie-Gruppe mit endlich vielen Komponenten. J ist kompakt und bei geeigneter Koordinatenwahl eine lineare orthogonale Gruppe, die sich auf der Abstandssphäre transitiv benimmt.*

2.3. Sei F' eine offene abgeschlossene Untergruppe von F. Sei $J' = J \cap F'$. Die kanonische Beziehung zwischen den Nebengruppenräumen F/J und F'/J' ist dann eine Homöomorphie.

Beweis. Jede Nebengruppe von F' in F enthält Elemente von J, da sonst $R = F/J$ nicht zusammenhinge. Wir bilden F/J auf F'/J' ab:

$$fJ \to fJ \cap F'.$$

$fJ \cap F' = f(J \cap f^{-1} F')$ ist nicht leer. Ist $f_1 \in fJ \cap F'$ und $f_2 \in J'$, so ist $f_1 f_2 \in fJ \cap F'$. Ist $f_i \in fJ \cap F'$, so ist $f_1^{-1} f_2 \in J'$. Also ist $fJ \cap F'$ eine Nebengruppe von J' in F'. Die Beziehung ist eineindeutig, denn aus $f_1 J \cap F' = f_2 J \cap F'$ folgt $f_1^{-1} f_2 \in J \cap F'$. Sie bezieht F/J *auf* F'/J', denn für ein gegebenes fJ' mit $f \in F'$ gilt: $fJ \to fJ \cap F' = f(J \cap F') = fJ'$. Der topologische Charakter ist klar.

2.4. Der Bahnbegriff ändert sich, wenn man (unter den Voraussetzungen von 2.3) von F/J nach F'/J' übergeht. Wir zeigen: Eine J-Bahn in F/J zerfällt in *endlich* viel J'-Bahnen in F'/J'.

J ist nämlich kompakt, also ist auch J/J' kompakt, aber außerdem, da J' offen in J ist, diskret. Also ist J/J' endlich. Die J-Bahn zerfällt in höchstens soviel J'-Bahnen wie J in Nebenklassen nach J'.

2.5. F_0 sei die Einskomponente von F. Dann ist F_0 Normalteiler von F.

F_0 ist zusammenhängend und im Kleinen kompakt, besitzt also beliebig kleine kompakte Untergruppen F_1, so daß F_0/F_1 Liesch ist[10]).

Wir zeigen, daß der Normalisator F' von F_1 in F offen (und abgeschlossen) in F ist:

Mit F_1 ist auch fF_1f^{-1} Normalteiler in F_0, also auch $F_1 \cdot fF_1f^{-1}$. Nun ist $F_1 \cdot fF_1f^{-1}/F_1$ Normalteiler der Lie-Gruppe F_0/F_1 und andererseits in einer vorgeschriebenen Umgebung der Eins enthalten, wenn f genügend nahe der Eins gewählt wird. Das ist aber nur möglich, wenn $F_1 \cdot fF_1f^{-1} = F_1$ ist. Das bedeutet wiederum $fF_1f^{-1} = F_1$ für *alle* f in der Nähe der Eins. Der Normalisator von F_1 ist also offen.

2.6. Die Forderung (Z) kann auch geschrieben werden:

$$JfJ \text{ zerlegt } F \text{ für gewisses } f \in F.$$

Wir wählen nun F_1 nach 2.5 so klein, daß auch noch gilt:

$$JfF_1J \text{ zerlegt } F.$$

Wir definieren F' als Normalisator von F_1 in F und können nach 2.5 (Ende) nun 2.3 (auf F_1 statt F') anwenden. Von f darf man nach 2.3 voraussetzen: $f \in F'$. Wir unterwerfen JfF_1J/J der kanonischen Homöomorphie aus 2.3. Dann ergibt sich:

$$JfF_1J' = JfF_1J \cap F' \quad \text{zerlegt } F'.$$

Nach 2.4 gibt es endlich viele $f_i \in F'$, so daß

$$JfF_1J' = \bigcup_i J'f_iF_1J'.$$

Da F_1 Normalteiler in F' ist, darf man das auch schreiben:

$$(*) \qquad \bigcup_i F_1J'f_iF_1J' \quad \text{zerlegt } F'.$$

Mit der kanonischen Abbildung

$$\varphi: \quad F' \to F'/F_1 = F'',$$
$$R = F'/J' \to F'/F_1J' = R'',$$
$$f_i \to f_i'',$$
$$F_1J' \to F_1J'/F_1 = J''$$

erhält man eine gemischte Lie-Gruppe F'', die transitiv wirkt in R'' und da eine Standuntergruppe J'' besitzt. $(*)$ besagt:

Z': R'' wird durch eine gewisse endliche Vereinigung von Bahnen zerlegt.

[10]) YAMABE, H.: Ann. of Math. (2) **58**, 351—365 (1953).

Da R'' eine (n-dimensionale) analytische Mannigfaltigkeit ist und jede Bahn in R'' als Vereinigung analytischer Mannigfaltigkeiten liegt, ist mindestens eine Bahn (nach Z') $(n-1)$-dimensional.

J'' ist kompakt und besitzt daher nur endlich viel Komponenten; die Einskomponente heiße J_0''. Da wegen des Zusammenhanges von R'' jede Komponente von F'' Punkte von J'' enthält, besitzt F'' höchstens soviel Komponenten wie J'', also nur endlich viele. Die Einskomponente von F'' heiße F_0''.

2.7. Die $j \in J''$ induzieren im Infinitesimalring $\boldsymbol{F}_0''$ lineare Abbildungen

$$\psi_j : \boldsymbol{f} \to j \boldsymbol{f} j^{-1},$$

die J'' darstellen. Wegen der Kompaktheit von J'' sind sie nach geeigneter Abstandsdefinition orthogonal. Die ψ_j lassen $\boldsymbol{J}_0''$ und sein orthogonales Komplement $\boldsymbol{K}$ invariant. Man darf in der Nähe von x_0 nun R'' mit $\exp \boldsymbol{K}$ identifizieren: man ordne dem Punkte fJ'' von R den (dort eindeutig bestimmten) Schnittpunkt von $\exp \boldsymbol{K}$ mit der Nebenklasse fJ'' zu.

Die Punkte von R'' stellen sich dann im Kleinen in der Form

$$\exp \boldsymbol{k} \cdot J'' \qquad (\boldsymbol{k} \in \boldsymbol{K})$$

dar. Anwendung eines $j \in J''$ ergibt

$$\exp \boldsymbol{k} \cdot J'' \to j \exp \boldsymbol{k} \cdot J'' = j \exp \boldsymbol{k} j^{-1} J'' = \exp j \boldsymbol{k} j^{-1} \cdot J''.$$

j transformiert R dann gemäß ψ_j.

J'' benimmt sich also in der Nähe von x_0 wie eine orthogonale Gruppe. Da eine Bahn von J'' $(n-1)$-dimensional ist, sind alle seine Bahnen im allgemeinen mindestens $(n-1)$-dimensional. Hieraus folgt, daß J'' auf der Abstandssphäre transitiv ist.

2.8. Wir kehren zu F zurück. Wir hatten $F_0'' = F_0 / F_1$. Ist F_0 nicht Lie-Gruppe, so gibt es eine echt absteigende Folge F_n von Normalteilern von F_0 so, daß die F_0/F_n Lie-Gruppen sind, die die in Satz I ausgedrückten Eigenschaften besitzen. Es gibt dann eine echte unendliche Folge stetiger kanonischer Homomorphismen

$$F_0/F_1 \leftarrow F_0/F_2 \leftarrow \cdots.$$

Nach der Klassifikation der dem Satz I entspringenden Lie-Gruppen zeigt sich aber (s. 8), daß solche Folgen höchstens die Länge 5 haben können. F_0 muß daher eine Lie-Gruppe sein und F eine gemischte Lie-Gruppe, die dem Satz I genügt.

3. Die zulässigen reinen Standuntergruppen

3.1. Wir behalten die Bezeichnungen bei, lassen aber die doppelten Akzente weg.

Mit J besitzt auch J_0 die Eigenschaften: in sich kompakt, linear, orthogonal, transitiv auf der Sphäre.

Wir behandeln vorläufig nur J_0, lassen aber den Index weg, setzen also Zusammenhang voraus.

J wirkt in einem linearen Raum $\boldsymbol{K}$.

3.2. J ist sicher (reell) irreduzibel, also ist $\boldsymbol{J}$ halbeinfach oder direkte Summe eines halbeinfachen und eines eindimensionalen Lie-Ringes. Die zweite Möglichkeit besteht nur, wenn $\boldsymbol{J}$ komplex reduzibel ist.

Da J orthogonal ist, sind alle seine Wurzelformen und Gewichte rein imaginär (bezogen auf eine reelle maximal Abelsche reguläre Untergruppe H). Mit jeder Wurzelform kommt also auch die entgegengesetzte vor und mit jedem Gewicht das entgegengesetzte. Wenn x ein Vektor von $\boldsymbol{K}$ vom Gewicht g ist, so ist $\bar{x}$ einer vom Gewicht $-g$. Das zur Wurzel α gehörige Element von $\boldsymbol{J}$ wird wie üblich mit e_α bezeichnet.

3.3. Seien x und y Vektoren von $\boldsymbol{K}$ vom Gewicht $g \neq 0$ und $g' \neq 0$. Wegen der Transitivität von J auf der Sphäre muß es ein $j \in \boldsymbol{J}$ geben mit

$$j(x + \bar{x}) \equiv y + \bar{y} \mod x + \bar{x}.$$

Sei

$$j = h + \Sigma(\tau_\alpha e_\alpha + \bar{\tau}_\alpha \bar{e}_\alpha) \quad \text{mit } h \in \boldsymbol{H}.$$

Also

$$\left(h + \Sigma\left(\tau_\alpha e_\alpha + \tau_\alpha \bar{e}_\alpha\right)\right)(x + \bar{x}) \equiv y + \bar{y} \mod x + \bar{x}.$$

Links steht eine Kombination von Vektoren der Gewichte $\pm g$, $\pm(g \pm \alpha)$ (α = variable Wurzelform), rechts eine von Vektoren der Gewichte $\pm g'$. Also:

3.3.1. Für je zwei Gewichte $g \neq 0$, $g' \neq 0$ mit $g \neq \pm g'$ gilt:

$$g + g' \quad \text{oder} \quad g - g' \quad \text{ist Wurzelform.}$$

Sei nun $g = g' \neq 0$ (aber $x, y, \bar{x}, \bar{y}$ unabhängig). Dann ergibt sich analog:

3.3.2. Ist das Gewicht $g \neq 0$ von einer Vielfachheit > 1, so ist diese Vielfachheit $= 2$ und $2g$ Wurzelform.

Schließlich hat man:

3.3.3. Das Gewicht 0 kann höchstens die Vielfachheit 1 besitzen, in dem Falle ist jedes Gewicht Wurzelform.

3.4. Ist $\boldsymbol{J}$ nicht-trivial direkte Summe zweier Lie-Ringe $\boldsymbol{L}_1, \boldsymbol{L}_2$, so ist 0 nicht Gewicht und einer von beiden hat (in der durch $\boldsymbol{J}$ induzierten Darstellung) höchstens zwei verschiedene Gewichte.

Beweis. Der maximal abelsche reguläre Teilring von $\boldsymbol{L}_i$ heiße $\boldsymbol{H}_i$. Jede Wurzelform verschwindet entweder auf $\boldsymbol{H}_1$ oder auf $\boldsymbol{H}_2$. Ist 0 Wurzelform, so verschwindet nach 3.3.3 jedes Gewicht auf $\boldsymbol{H}_1$ oder auf $\boldsymbol{H}_2$. Wäre g_i ($i = 1, 2$) ein Gewicht, das auf $\boldsymbol{H}_i$ nicht verschwände, so käme man mit dem Gewicht $g_1 \pm g_2$ (s. 3.3.1) zu einem Widerspruch. Also verschwinden alle Gewichte etwa auf $\boldsymbol{H}_2$. Also $\boldsymbol{L}_2 = 0$ im Widerspruch zur Voraussetzung.

Jedes Gewicht g von $\boldsymbol{J}$ ist zu schreiben als Summe von Linearformen

$$g = g_1 + g_2,$$

wo g_1 auf $\boldsymbol{H}_2$ verschwindet und g_2 auf $\boldsymbol{H}_1$. Mit einem weiteren Gewicht

$$g' = g_1' + g_2' \quad \text{mit} \quad g_1' \neq \pm g_1$$

erhalten wir nach 3.3.1 eine Wurzelform $g \pm g'$, die auf $\boldsymbol{H}_1$ oder auf $\boldsymbol{H}_2$ verschwinden muß. Alle g' haben also die Form

$$\pm g_1 + g_2' \quad \text{oder} \quad g_1' \pm g_2.$$

Diese beiden widersprechen einander nach dem soeben Bewiesenen, wenn nicht $g_1 = \pm g_1'$ oder $g_2 = \pm g_2'$. Hieraus folgt die Behauptung.

3.5. Wir nehmen im folgenden an, daß $\boldsymbol{L}_2$ nur zwei verschiedene Gewichte besitzt. $\boldsymbol{L}_2$ ist dann vom Rang 0 oder vom Rang 1. Im zweiten Fall ist $\boldsymbol{K}$ komplexe Summe von Ebenen, in denen $\boldsymbol{L}_2$ wie $\pi_1(\boldsymbol{A}_1)$ operiert[11]).

Jedenfalls erzeugen die Vektoren von Gewichten $g_1 + g_2$ bzw. $g_1 - g_2$ einen Unterraum $\boldsymbol{K}_+$ bzw. $\boldsymbol{K}_-$ von $\boldsymbol{K}$. Ist $x \in \boldsymbol{K}_+$ ein Vektor vom Gewicht g, so ist $\bar{x}$ vom Gewicht $-g$, also in $\boldsymbol{K}_-$. Daher sind $\boldsymbol{K}_+$ und $\boldsymbol{K}_-$ konjugiert komplex. Ihre (direkte) Summe ist $\boldsymbol{K}$.

$\boldsymbol{K}_+$ und $\boldsymbol{K}_-$ sind invariant gegenüber $\boldsymbol{L}_1$, da die Basis-Operationen von $\boldsymbol{L}_1$ einen Vektor vom Gewicht $g_1 + g_2$ in einen von einem Gewicht $g_1 + g_2$ oder $g_1 + \alpha + g_2$ (α Wurzelform von $\boldsymbol{L}_1$) überführen.

Ist $\boldsymbol{L}_2$ vom Range 1, so vertauscht die Spiegelung an der Wurzelebene $g_2 = 0$ ($2g_2$ ist Wurzel von $\boldsymbol{L}_2$) die $g_1 \pm g_2$ untereinander. $\boldsymbol{K}_+$ und $\boldsymbol{K}_-$ werden durch $\boldsymbol{L}_1$ dann äquivalent transformiert.

In keinem Fall kann $\boldsymbol{L}_1$ (in $\boldsymbol{K}_+$) einer reellen Darstellung äquivalent sein. Sonst wäre nämlich die Abbildung $x \to \bar{x}$ ($x \in \boldsymbol{K}$) vertauschbar mit den Elementen von $\boldsymbol{L}_1$; die Teilräume $\boldsymbol{K}'$ bzw. $\boldsymbol{K}''$ der $x + \bar{x}$ bzw. $i(x - \bar{x})$ mit $x \in \boldsymbol{K}_+$ wären ebenfalls invariant bei $\boldsymbol{L}_1$ und obendrein reell; die Darstellungen von $\boldsymbol{L}_1$ in $\boldsymbol{K}'$ und $\boldsymbol{K}''$ wären reell äquivalent; $\boldsymbol{L}_2$, das mit $\boldsymbol{L}_1$ vertauschbar und orthogonal ist, wäre vom Range 0; die Elemente von $\boldsymbol{L}$ hätten die Form

$$\begin{pmatrix} M & \varrho \cdot 1 \\ -\varrho \cdot 1 & M \end{pmatrix},$$

und $\boldsymbol{J}$ könnte nicht transitiv auf der Sphäre sein.

3.6. Für die Gewichte von $\boldsymbol{L}_1$ (d.h. die Gewichte von $\boldsymbol{J}$, auf $\boldsymbol{H}_1$ betrachtet) gelten wieder die Behauptungen 3.3.1–3, jedoch mit einigen Modifikationen:

3.6.1. Für je zwei Gewichte $g_1 \neq 0$, $g_1' \neq 0$ mit $g_1 \neq \pm g_1'$ gilt:

$$g_1 + g_1' \quad \text{oder} \quad g_1 - g_1' \quad \text{ist Wurzelform.}$$

3.6.2. Alle Gewichte $g_1 \neq 0$ sind einfach.

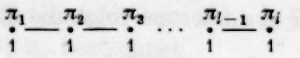
3.6.3. Das Gewicht 0 kann höchstens die Vielfachheit 2 besitzen.

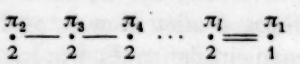
[11]) Siehe Vorbemerkungen 4.

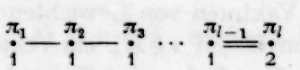
Hier ist 3.6.1 klar. — Ist g_1 mehrfach, so gibt es ein mehrfaches Gewicht g_1+g_2 von $\boldsymbol{J}$; nach 3.3.2 ist dann $2g_1+2g_2$ Wurzelform, also $g_1=0$ und g_2 von der Vielfachheit 2. Hieraus folgen 3.6.2—3.

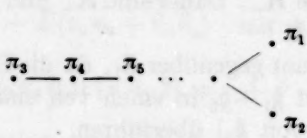
3.7. Wir belegen $\boldsymbol{J}$, wenn es einfach ist, und $\boldsymbol{L}_1$ (im Raum $\boldsymbol{K}_1$), wenn $\boldsymbol{J}$ direkte Summe ist, mit dem gemeinsamen Namen $\boldsymbol{L}$. Wir untersuchen vorläufig $\boldsymbol{L}$. Es ist klar, daß $\boldsymbol{L}$ einfach ist — eventuell auch $=\boldsymbol{O}$ oder eindimensional, aber diese Fälle lassen wir vorläufig weg.

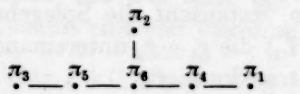
3.8. Bei den einfachen Gruppen bedienen wir uns der folgenden Schemata:

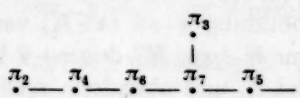
$\boldsymbol{A}_l$:  , $l \geq 1$.

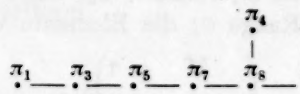
$\boldsymbol{B}_l$:  , $l \geq 2$.

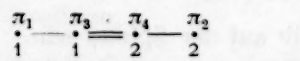
$\boldsymbol{C}_l$:  , $l \geq 3$.

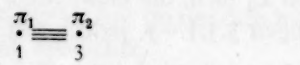
$\boldsymbol{D}_l$:  , $l \geq 4$.

$\boldsymbol{E}_6$: 

$\boldsymbol{E}_7$: 

$\boldsymbol{E}_8$: 

$\boldsymbol{F}_4$: 

$\boldsymbol{G}_2$: 

Die irreduziblen Wurzelformen ϱ_i sind als dicke Punkte angedeutet. Bei einem ϱ_i ist das zugehörige fundamentale dominante Gewicht π_i angegeben, d.h.

$$2\frac{(\pi_j \cdot \varrho_i)}{(\varrho_i \cdot \varrho_i)} = \begin{cases} 1 \text{ für } j = i \\ 0 \text{ für } j \neq i. \end{cases}$$

Wir notieren noch die dominanten Wurzelformen $\neq 0$:

A_l: $\pi_1 + \pi_l$ (für $l = 1 : \pi_1$).
B_l: π_2, π_3 (für $l = 2 : \pi_2, 2\pi_1$).
C_l: $2\pi_1, \pi_2$.
D_l: π_4.
E_6: π_2.
E_7: π_3.
E_8: π_1.
F_4: π_1, π_2.
G_2: π_1, π_2.

3.9. Sei g ein dominantes Gewicht von $\boldsymbol{L}$ mit der Eigenschaft:

$$2\frac{(g,\varrho)}{(\varrho,\varrho)} \geq 2$$

für eine gewisse irreduzible Wurzelform ϱ. Dann ist $g - 2\varrho$ Gewicht, also ist $2g - 2\varrho = \alpha$ nach 3.3.1 und 3.6.1 Wurzelform (denn 2ϱ ist sicher keine). $\alpha = 0$ bedeutet, daß das dominante Gewicht g auch irreduzible Wurzelform ist, also Rang $= 1$. Im Falle $\alpha \neq 0$ zeigt sich, daß α dominante Wurzelform ist. Denn $2\frac{(\alpha,\varrho)}{(\varrho,\varrho)} \geq 0$ und für jeden Nachbarn ϱ' von ϱ:

$$2\frac{(\alpha,\varrho')}{(\varrho',\varrho')} > 0.$$

Weiter ist

$$2\frac{(\alpha,\beta)}{(\beta,\beta)} \text{ gerade}$$

für jede Wurzelform β. Das ist nur möglich beim Rang $= 1$ oder bei $\alpha = 2\pi_1(\boldsymbol{B}_2)$ und $\alpha = 2\pi_1(\boldsymbol{C}_l)$ (s. 3.8). ϱ darf nur einen Nachbarn haben [außer für $\pi_2(\boldsymbol{A}_3)$, aber das ist schon weggefallen]. Im Falle $\alpha = 2\pi_1(\boldsymbol{C}_l) = 2\omega_1$, wird $\varrho = \omega_1 - \omega_2$, $g = 2\omega_1 - \omega_2$, was nicht dominant ist. Im Falle $\alpha = 2\pi_1(\boldsymbol{B}_2) = \omega_1 + \omega_2$, $\varrho = \omega_1 - \omega_2$ oder $= \omega_2$, $g = \frac{3}{2}\omega_1 - \frac{1}{2}\omega_2$ oder $\frac{1}{2}\omega_1 + \frac{3}{2}\omega_2$, die nicht dominant sind.

Nur der Rang 1 bleibt übrig. Mit 3.3.1 und 3.6.1 verträglich sind $g = p\pi_1(\boldsymbol{A}_1)$ $p = 1, 2, 3$. Die irreduzible Darstellung mit höchstem Gewicht $3\pi_1(\boldsymbol{A}_1)$ ist 4-dimensional und wesentlich nicht reell[12]). $\boldsymbol{J}$ ist dann höchstens $(3+3)$-dimensional und kann auf der Sphäre im $2 \cdot 4$-dimensionalen Raum nicht transitiv sein. $3\pi_1(\boldsymbol{A}_1)$ fällt also auch weg.

Nun bleiben $\pi_1(\boldsymbol{A}_1)$, $2\pi_1(\boldsymbol{A}_1)$ übrig.

3.10. Wir dürfen jetzt voraussetzen:

$$2\frac{(g,\varrho)}{(\varrho,\varrho)} \leq 1$$

für jedes dominante Gewicht g und jede irreduzible Wurzelform ϱ.

[12]) Die Realitätsfragen bei Darstellungen werden wir an anderer Stelle allgemein behandeln. — Alle $p\pi_1(\boldsymbol{A}_1)$ mit ungeradem p sind nichtreell.

Sei

$$(*) \qquad 2\frac{(g,\varrho)}{(\varrho,\varrho)} = 1 \quad \text{und} \quad 2\frac{(g,\varrho')}{(\varrho',\varrho')} = 1$$

für zwei irreduzible Wurzelformen ϱ, ϱ'. Dann sind $g-\varrho$ und $g-\varrho'$ Gewichte. $\varrho-\varrho'$ ist nicht Wurzelform, also (nach 3.3.1 und 3.6.1) $\alpha = 2g-\varrho-\varrho'$ wohl. $\alpha = 0$ bedeutet $g = \frac{1}{2}(\varrho+\varrho')$; das kann aber nur beim Rang 2 dominant sein; man sieht leicht, daß das unmöglich ist.

α ist wieder dominant und

$$2\frac{(\alpha,\varrho'')}{(\varrho'',\varrho'')} > 0$$

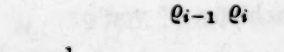
für alle Nachbarn ϱ'' von ϱ oder ϱ'. Nur $\pi_3(\boldsymbol{B}_3)$, $\pi_2(\boldsymbol{C}_3)$ lassen das zu, erweisen sich aber auch als unmöglich. Daher ist (*) unzulässig und g notwendig fundamental.

Im folgenden ist ϱ_i die irreduzible Wurzelform, für die

$$2\frac{(g,\varrho_i)}{(\varrho_i,\varrho_i)} = 1$$

gilt. Wir unterscheiden verschiedene Fälle.

3.11.



Hier sind $g-\varrho_i$ und $g-\varrho_{i-1}-\varrho_i-\varrho_{i+1}$ Gewichte. $\varrho_{i-1}+\varrho_{i+1}$ ist nicht Wurzelform. Daher ist (3.3.1 und 3.6.1) $\alpha = 2g-\varrho_{i-1}-2\varrho_i-\varrho_{i+1}$ wohl Wurzelform, und zwar wieder 0 oder dominant.

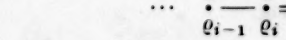
$$2\frac{(\alpha,\varrho)}{(\varrho,\varrho)} > 0 \quad \text{für} \quad \varrho = \varrho_{i-2}, \varrho_{i+2},$$
$$= 0 \quad \text{sonst.}$$

Mit 3.8 ist das nur verträglich, wenn entweder g an einem der Schemaenden steht oder an einem Kreuzpunkt oder eines der Gewichte

$$\pi_2(\boldsymbol{A}_3),\ \pi_3(\boldsymbol{A}_5),\ \pi_l(\boldsymbol{B}_l),\ \pi_{l-1}(\boldsymbol{C}_l),\ \pi_3(\boldsymbol{F}_4),\ \pi_4(\boldsymbol{F}_4)$$

ist. Nun ist die Darstellung zu $\pi_3(\boldsymbol{A}_5)$ wesentlich komplex[13]; die reelle Darstellung geschieht in $2\cdot 20$ Dimensionen; die Gruppe hat höchstens $35+3$ Dimensionen; es herrscht also keine Transitivität auf der Sphäre.

3.12.



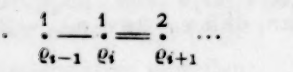
Wieder ist $\alpha = 2g-\varrho_{i-1}-2\varrho_i-\varrho_{i+1} = 0$ oder dominante Wurzelform. Hier wird

$$2\frac{(\alpha,\varrho)}{(\varrho,\varrho)} > 0 \quad \text{auch noch für} \quad \varrho = \varrho_{i+1}.$$

Hiermit fallen die $\pi_l(\boldsymbol{B}_l)$ fort, außer $\pi_2(\boldsymbol{B}_2)$. Weiter fällt $\pi_4(\boldsymbol{F}_4)$ fort.

[13]) Siehe Fußnote 12. — Die Darstellungen $\pi_k(\boldsymbol{A}_{2k-1})$ sind nur für gerades k reell.

3.13.

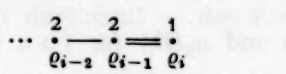


$\alpha = 2g - \varrho_{i-1} - 2\varrho_i - \varrho_{i+1} = 0$ oder dominant.

$$2\frac{(\alpha, \varrho)}{(\varrho, \varrho)}$$

benimmt sich wie in 3.11; aber auch noch > 0 für $\varrho = \varrho_i$. Demnach fallen die $\pi_{l-1}(C_l)$ fort, außer $\pi_2(C_3)$. Im letzten Fall besitzt die Darstellung zwei Vektoren vom Gewicht 0, was nach 3.3.2 unmöglich ist für $\boldsymbol{L} = \boldsymbol{J}$; für $\boldsymbol{L} = \boldsymbol{L}_1$ ist die Dimension der Darstellung wieder zu hoch. Also fällt auch $\pi_2(C_3)$ fort. Ferner fällt noch $\pi_3(F_4)$ fort.

3.14.

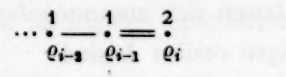


Hier sind ϱ_i und $\varrho_{i-2} + 2\varrho_{i-1} + 2\varrho_i$ Wurzelformen, ihre Differenz aber nicht. $\alpha = (g - \varrho_i) + (g - \varrho_{i-2} - 2\varrho_{i-1} - 2\varrho_i) = 0$ oder dominante Wurzelform

$$2\frac{(g, \varrho)}{(\varrho, \varrho)} > 0 \quad \text{für} \quad \varrho = \varrho_{i-3},$$
$$= 0 \quad \text{sonst.}$$

Hierdurch fällt $\pi_1(B_l)$ fort für $l > 5$. Für $l = 5$ ist die irreduzible Darstellung nicht reell[14]; die reelle ist 64-dimensional, die Gruppe ist höchstens $(55+3)$-dimensional. Also ist auch $\pi_1(B_5)$ unmöglich.

3.15.

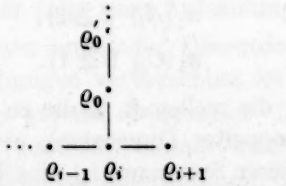


Wir nehmen die Wurzelformen ϱ_i, $\varrho_i + 2\varrho_{i-1} + \varrho_{i-2}$. Dann gilt für $\alpha = 2g - 2\varrho_i - 2\varrho_{i-1} - \varrho_{i-2}$

$$2\frac{(\alpha, \varrho)}{(\varrho, \varrho)} \begin{cases} > 0 & \text{für} \quad \varrho = \varrho_{i-1},\ \varrho_{i-3},\ \varrho_{i+1} \\ = 0 & \text{sonst.} \end{cases}$$

Damit fallen die $\pi_l(C_l)$ fort außer für $l = 3$. Auch $\pi_3(C_3)$ hat wieder keine irreduzible reelle Darstellung[15]; die Dimension der Darstellung wird $2 \cdot 20$, die der Gruppe höchstens $21 + 3$; also keine Transitivität. — Auch $\pi_2(F_4)$ fällt fort.

3.16.



[14]) Siehe Fußnote 12. — $\pi_1(B_l)$ ist reell für $l = 0, 3 \bmod 4$.

[15]) Siehe Fußnote 12. — Die $\pi_l(C_l)$ geben nur für gerades l reelle irreduzible Darstellungen.

Wie in 3.11 sieht man, daß $\alpha = 2g - \varrho_{i-1} - 2\varrho_i - \varrho_{i+1} = 0$ oder dominante Wurzelform ist mit

$$2\frac{(\alpha, \varrho)}{(\varrho, \varrho)} > 0 \quad \text{für} \quad \varrho = \varrho_{i-2}, \varrho_{i+2}, \varrho_0.$$

Alle $\pi_l(\boldsymbol{D}_l)$ und $\pi_8(\boldsymbol{E}_8)$ fallen darum fort. Ersetzt man ϱ_{i+1} durch ϱ_0, so sieht man, daß auch $\pi_6(\boldsymbol{E}_6)$ und $\pi_7(\boldsymbol{E}_7)$ fortfallen.

Ferner sind $\varrho_1 + \cdots + \varrho_i + \varrho_{i+1}$ und $\varrho_1 + \cdots + \varrho_i + \varrho_0$ Wurzelformen, aber ihre Differenz nicht. $\alpha = 2g = 2\varrho_1 - \cdots - 2\varrho_i - \varrho_{i+1} - \varrho_0$ ist 0 oder dominante Wurzelform mit

$$2\frac{(\alpha, \varrho)}{(\varrho, \varrho)} \begin{cases} > 0 & \text{für} \quad \varrho = \varrho_{i+2}, \varrho_0' \\ = 0 & \text{sonst.} \end{cases}$$

Hierdurch fallen $\pi_1(\boldsymbol{D}_l)$ und $\pi_2(\boldsymbol{D}_l)$ für $l > 4$ fort und alle $\pi_i(\boldsymbol{E}_j)$ in Endpunkten des Schemas.

3.17. $\pi_1(\boldsymbol{F}_4)$ besitzt zwei verschwindende Gewichte und ist reell. Also fällt es nach 3.3.2 fort. Dasselbe gilt für $\pi_2(\boldsymbol{G}_1)$.

Es ist zweckmäßig, bei $\boldsymbol{C}_l$ die Indizes 1, 2 zuzulassen, $\pi_1(\boldsymbol{B}_2)$ durch $\pi_1(\boldsymbol{C}_2)$ und $\pi_1(\boldsymbol{A}_1)$ durch $\pi_1(\boldsymbol{C}_1)$ zu ersetzen. Ebenso lasse man $\boldsymbol{D}_3$ zu und interpretiere man $\pi_2(\boldsymbol{A}_3)$ als $\pi_3(\boldsymbol{D}_3)$. Schließlich lasse man $\boldsymbol{B}_1$ zu und interpretiere man $2\pi_1(\boldsymbol{A}_1)$ als $2\pi_1(\boldsymbol{B}_1)$.

$\pi_1(\boldsymbol{A}_l)$ und $\pi_l(\boldsymbol{A}_l)$ sind als Gruppen identisch. Wir können uns auf eine beschränken. Dasselbe gilt für $\pi_1(\boldsymbol{D}_4)$, $\pi_2(\boldsymbol{D}_4)$, $\pi_3(\boldsymbol{D}_4)$.

3.18. Die Resultate lassen sich zusammenfassen in

Satz II. Die zulässigen reellen $\boldsymbol{L}$ sind

$\boldsymbol{O}$ (Nullabbildung im 1-dimensionalen $\boldsymbol{K}$),
$\pi_2(\boldsymbol{B}_l)$ $(l \geq 2)$, $2\pi_1(\boldsymbol{B}_1)$,
$\pi_1(\boldsymbol{B}_3)$, $\pi_1(\boldsymbol{B}_4)$,
$\pi_3(\boldsymbol{D}_l)$ $(l \geq 3)$,
$\pi_1(\boldsymbol{G}_2)$.

Die zulässigen komplexen $\boldsymbol{L}$ sind

$\pi(\boldsymbol{A}_0)$,
$\pi_1(\boldsymbol{A}_l)$ $(l \geq 2)$,
$\pi_1(\boldsymbol{C}_l)$ $(l \geq 1)$.

Die zulässigen $\boldsymbol{J}$ sind: 1. die reellen $\boldsymbol{L}$, 2. die zu den komplexen $\boldsymbol{L}$ gehörigen reellen Darstellungen (doppelter Dimension), eventuell erweitert mit einem direkten Summand — dieser Summand ist eine Darstellung von $\boldsymbol{A}_0$ im Falle $\pi_1(\boldsymbol{A}_l)$ $(l \geq 2)$ und eine Darstellung von $\boldsymbol{A}_0$ oder $\boldsymbol{A}_1$ im Falle $\pi_1(\boldsymbol{C}_l)$ und im Falle $\pi_1(\boldsymbol{A}_1)$.

Die zulässigen reinen Standuntergruppen sind hiermit bestimmt.

Wir bezeichnen diese Gruppen mit $\nu(\boldsymbol{O}), \nu(\boldsymbol{B}_1), \ldots, \nu(\boldsymbol{C}_l+\boldsymbol{A}_1)$. Die $\pi_1(\boldsymbol{B}_3), \pi_1(\boldsymbol{B}_4)$ unterscheiden wir als $\nu'(\boldsymbol{B}_3), \nu'(\boldsymbol{B}_4)$.

Wir können sie auch folgendermaßen einteilen:

1. die Drehgruppen $\dot{\mathrm{G}}_0\mathrm{UVR}(n, \mathrm{re})$ — das sind die $\nu(\boldsymbol{B}_l), \nu(\boldsymbol{D}_l), \nu(\boldsymbol{O}), \nu(\boldsymbol{A}_0), \nu(\boldsymbol{C}_1+\boldsymbol{A}_1)$.

2. die unitären Gruppen (bzw. mit der Determinante 1) $\mathrm{GUVR}(n, \mathrm{ko})$ bzw. $\mathrm{G}_0\mathrm{UVR}(n, \mathrm{ko})$ — das sind die $\nu(\boldsymbol{A}_l+\boldsymbol{A}_0)$ bzw. die $\nu(\boldsymbol{A}_l)$ $(l \geqq 2)$;

3. die quaternion-unitären Gruppen $\mathrm{GUVR}(n, \mathrm{qu})$ der Transformationen

$$x_i' = \sum a_{ij}\, x_j\, b,$$

wo $\sum_j a_{ij}\bar{a}_{ik} = \delta_{ik}$ ist und $b=1$ ist oder eine eindimensionale Gruppe oder alle Quaternionen mit $|b|=1$ durchläuft — das sind die $\nu(\boldsymbol{C}_l), \nu(\boldsymbol{C}_l+\boldsymbol{A}_0), \nu(\boldsymbol{C}_l+\boldsymbol{A}_1)$;

4. die (reine) Gruppe der Automorphismen der Oktavenalgebra (betrachtet für die Oktaven mit verschwindendem Realteil) — das ist die $\nu(\boldsymbol{G}_2)$;

5. die von den infinitesimalen Transformationen

$$x \to \sum_1^7 a_{ij}\, e_i\, (e_j\, x)$$

erzeugte Gruppe [x Oktave, e_i $(i=0, \ldots, 7)$ Oktavenbasiselemente] — das ist die $\nu'(\boldsymbol{B}_3)$;

6. die oktaven-unitäre Gruppe (in der Oktavenebene) $\mathrm{GUVR}(2, \mathrm{ok})$ — das ist die $\nu'(\boldsymbol{B}_4)$.

Die beiden letzten sind die niedrigsten Spindarstellungen der $\boldsymbol{B}_3$ und $\boldsymbol{B}_4$.

4. Die zulässigen gemischten Standuntergruppen

4.1. Die Einskomponente J_0 von J ist aus Satz II zu entnehmen. J_0 ist Normalteiler von J. Die Elemente von J induzieren Automorphismen in J_0. Um J zu kennen, braucht man nur in jeder Nebenklasse von J_0 einen Repräsentanten. Hierzu muß man in der Automorphismengruppe von J_0 die Nebenklassen der inneren Automorphismengruppe kennen und weiter im Zentralisator von J_0 innerhalb der allgemeinen orthogonalen Gruppe die Nebenklassen nach dem Zentrum von J_0.

Die Automorphismenklassen sind wohlbekannt[16]), die Zentralisatorklassen sind leicht festzustellen. Hier folgt eine Aufzählung:

1. Drehgruppen in Räumen ungerader Dimension: keine äußeren Automorphismen. Mit allen Drehungen vertauschbar ist nur ± 1.

Drehgruppen in Räumen gerader Dimension: Die Spiegelung an einer Hyperebene repräsentiert die äußeren Automorphismen. Zentralisator trivial.

$\boldsymbol{D}_4$ besitzt wohl mehr äußere Automorphismen, die aber auf $\pi_3(\boldsymbol{D}_4)$ nicht eindeutig sind.

[16]) Cartan, E.: Oeuvres I, 555.

2. Unitäre Gruppen: Äußerer Automorphismus $x \to \bar{x}$, also realisierbar als lineare Abbildung in $\boldsymbol{K}$. Die mit $\nu(\boldsymbol{A}_l)$ elementweise vertauschbaren befinden sich in $\nu(\boldsymbol{A}_l + \boldsymbol{A}_0)$.

3. Quaternion-unitäre Gruppen: $\boldsymbol{C}_l$ besitzt keinen äußeren Automorphismus. Die mit $\nu(\boldsymbol{C}_l)$ elementweise vertauschbaren befinden sich in $\nu(\boldsymbol{C}_l + \boldsymbol{A}_1)$.

4. $\nu(\boldsymbol{G}_2)$: Kein äußerer Automorphismus. Mit allen vertauschbar ist nur ± 1.

5. $\nu'(\boldsymbol{B}_3)$: Kein äußerer Automorphismus. Zentralisator besteht aus ± 1, aber da $\nu'(\boldsymbol{B}_3)$ eine Spindarstellung ist und also ein nichttriviales Zentrum besitzt, befinden ± 1 sich in $\nu'(\boldsymbol{B}_3)$.

6. $\nu'(\boldsymbol{B}_4)$: wie $\nu'(\boldsymbol{B}_3)$.

5. Der Einbau der reinen Stänuntergruppen

5.1. Wir untersuchen nun, in welchen Gruppen $\boldsymbol{F}_0$ (Einskomponente von $\boldsymbol{F}$) die erhaltenen $\boldsymbol{J}_0$ liegen können, und auf welche Weise dies möglich ist. Wir lassen die Indizes 0 vorläufig weg.

Klar ist, daß gelten muß:

5.1.1. $\dim \boldsymbol{F} = \dim \boldsymbol{J} + \dim \boldsymbol{K}$.

5.1.2. J kann keinen Normalteiler von F enthalten, da die Darstellung in F/J sonst nicht treu wäre.

5.1.3. Bei der von J in F induzierten Gruppe von Automorphismen von F ist $\boldsymbol{J}$ invarianter Unterraum; mod $\boldsymbol{J}$ ist diese Darstellung im wesentlichen J selber, also reell irreduzibel (s. 2.6).

5.2. Aus 5.1.3 folgt: ist $\boldsymbol{F}_1$ (reelles) Ideal von $\boldsymbol{F}$, so ist $\boldsymbol{F}_1$ invarianter Unterraum (bei den von J induzierten Automorphismen von $\boldsymbol{F}$), also

$$\boldsymbol{F}_1 + \boldsymbol{J} \equiv (\boldsymbol{O}) \text{ oder } \equiv \boldsymbol{F} \mod \boldsymbol{J}.$$

Nach 5.1.2 ist $\boldsymbol{F}_1 < \boldsymbol{J}$ unmöglich, also

$$\boldsymbol{F}_1 + \boldsymbol{J} = \boldsymbol{F}.$$

Wir können uns beschränken auf die Annahmen: $\boldsymbol{F}_1$ ist halbeinfach, oder $\boldsymbol{F}_1$ ist abelsch.

5.3. Sei $\boldsymbol{F}_1$ halbeinfach. Dann gibt es eine direkte Summenzerlegung in Ideale

$$\boldsymbol{F} = \boldsymbol{F}_1 + \boldsymbol{F}_2.$$

Entsprechend 5.2 ist auch

$$\boldsymbol{F}_1 + \boldsymbol{J} = \boldsymbol{F}.$$

Wir zeigen:

$$\boldsymbol{F}_1 \cap \boldsymbol{J} = \boldsymbol{F}_2 \cap \boldsymbol{J} = (\boldsymbol{0}).$$

Sei $\boldsymbol{f} \in \boldsymbol{F}_2 \cap \boldsymbol{J}$ und $\boldsymbol{f}_2 \in \boldsymbol{F}_2$. Also

$$[\boldsymbol{f}, \boldsymbol{f}_2] \in \boldsymbol{F}_2.$$

Weiter ist f_2 darstellbar als

$$f_2 = f_1 + j \quad \text{mit} \quad f_1 \in F_1, \quad j \in J.$$

Also auch

$$[f, f_2] = [f, j] \in J.$$

Daher

$$[f, f_2] \in F_2 \cap J.$$

Also ist $F_2 \cap J$ Ideal von F_2, also auch von F. Nach 5.1.2 folgt hieraus $F_2 \cap J = (O)$. Ebenso ergibt sich $F_1 \cap J = (0)$.

Die Zerlegung von J nach F_1 und F_2,

$$j = f_1 + f_2 \quad (j \in J, \ f_i \in F_i)$$

bewirkt also einen Isomorphismus zwischen F_1 und F_2. Weiter ist $\dim J = \dim F_i = \dim K$. Ein Blick auf die Liste von Satz II lehrt, daß das nur möglich ist für $J = \nu(B_1)$. Dann wird F_i notwendig isomorph mit J.

Infolgedessen wird F (wenigstens im Kleinen) die Drehgruppe der 3-Sphäre.

5.4. Ist (statt 5.2—3) F_1 komplexes Ideal von F (aber halbeinfach), so ist F direkte Summe, etwa

$$F = F_1 + \overline{F}_1.$$

Ist $J \cap F_1 = J \cap \overline{F}_1 = (0)$, so ist $\dim J \leq \dim F_1$, also $\dim J \leq \dim K$, also $\dim J = \dim K$, also wieder $J = \nu(B_1)$. F_1 und $\overline{F}_1$ sind isomorph, und F wird (im Kleinen) die Gruppe des hyperbolischen 3-Raumes.

Wäre dagegen $J \cap F_1 \neq (0)$, so hätte J zwei fremde konjugierte isomorphe Ideale $J \cap F_1$ und $J \cap \overline{F}_2$, müßte also nach der Liste von Satz II entweder $\nu(A_0 + A_0)$ oder $\nu(A_1 + A_1)$ sein; dann wäre $\dim F = \dim J + \dim K = 4$ oder 10 und $\dim F_1 = 2$ oder 5, also F_1 nicht halbeinfach.

5.5. Sei (im Gegensatz zu 5.3) nun F_1 abelsch. Das orthogonale Komplement von $F_1 \cap J$ in F_1 sei F_2. Nun ist $F_2 = F_1 \cap K$, also

$$[J, F_2] \subset F_2.$$

$$[F_1, F_2] = (0),$$

also wegen $F = F_1 + J$ (s. 5.2) F_2 Ideal von F. Ferner $F = F_2 + J$, $F_2 \cap J = (0)$.

Ein komplexes F_1 liefert nichts Neues, da dann $F_1 + \overline{F}_1$ auch abelsches Ideal von F ist.

5.6. Aus 5.2—5 folgt nun

Satz III. *F ist entweder (komplex) einfach. Oder F zerfällt in J und ein abelsches Ideal. Oder F gehört zur elliptischen oder hyperbolischen Geometrie des 3-Raumes.*

Wir nennen allgemein die F mit abelschem Ideal *parabolisch*, die andern, je nachdem ob sie kompakt oder nichtkompakt sind, *elliptisch* und *hyperbolisch*.

5.7. Die parabolischen F bieten keine Schwierigkeiten mehr. Wir suchen die übrigen auf. Hierbei verwenden wir 5.1.1 und die Tatsache, daß der Rang von F mindesteus so groß ist wie der von J und höchstens um 1 größer (nämlich wenn das Gewicht 0 auftritt). In der folgenden Aufstellung findet man die Dimensionen einfacher Gruppen angegeben (erste Spalte) ferner die Summe der Dimensionen eines zugehörigen J (das die einfache Gruppe als Summand L_1 enthält) und K; oberhalb dieser Spalten findet man ein l oder $l+1$, je nachdem der Rang des eventuellen F gleich dem von L_1 ist oder um 1 größer. Aus dieser Tabelle suchen wir erst die Fälle aus, in denen sich wieder die Dimensionszahl einer einfachen Gruppe ergibt mit dem richtigen Rang.

A_l		l	$\leq l+1$	B_l		$\leq l+1$	C_l		l	$\leq l+1$		D_l		l
A_0	1	3		O	0	1								
A_1	3	7	8	B_1	3	6	C_1	3	7	8	10			
A_2	8	14	15	B_2	10	15	C_2	10	18	19	21			
A_3	15	23	24	B_3	21	28	C_3	21	33	34	36	D_3	15	21
A_4	24	34	35	B_4	36	45	C_4	36	52	53	55	D_4	28	36
A_5	35	47	48	B_5	55	66	C_5	55	75	76	78	D_5	45	55
A_6	48	62	63	B_6	78	91	C_6	78	102	103	105	D_6	66	78
A_7	63	79	80	B_7	105	120	C_7	105	133	134	136	D_7	91	105
⋮				⋮			⋮					⋮		

G_2 14 21 (rang 3).

E_6 78, E_7 133, E_8 248, F_4 52.

$\nu'(B_3)$ $21+8$, $\nu'(B_4)$ $36+16=52$.

Von den hiernach möglichen Paaren fallen noch folgende fort: $\nu(C_l+A_1)$ in B_{l+1} — die B_{l+1} ist reell linear darstellbar in $2l+1$ Dimensionen, die C_l erst in $4l$ Dimensionen; das ist ein Widerspruch für $l>1$. Der Fall $l=1$ ist die Einbettung von D_2 in B_3.

$\nu(D_l)$ in C_l — durch die Wurzeln und Gewichte von $J=\nu(D_l)$ ist F vollkommen bestimmt; als F kommt nur B_l in Frage.

$\nu(G_2)$ in C_3 — das ergäbe eine 6-dimensionale lineare Darstellung der G_2.

$\nu(C_4)$ in F_4 — $\nu(C_4)$ hat Wurzeln, die Doppelten von Gewichten sind; beide träten zusammen als Wurzeln von F auf.

$\nu(C_7)$ in E_7 — dieselben Gründe.

$\nu(C_5+A_1)$ in E_6 — hier hätte man eine Darstellung der E_6 als Transformationsgruppe einer 20-dimensionalen Mannigfaltigkeit, während so etwas erst in 26 Dimensionen möglich ist.

Satz IV. *Die möglichen Paare J, F im nichtparabolischen Fall sind:*

1. die n-dimensionale Drehgruppe in einer reellen Form der komplexen $(n+1)$-dimensionalen Drehgruppe.

2. $\nu(A_l+A_0)$ in einer reellen Form der komplexen A_{l+1} $(l\geq 1)$.

3. $\nu(C_l+A_1)$ in einer reellen Form der komplexen C_{l+1} $(l\geq 1)$.

4. $\nu(\boldsymbol{G}_2)$ *in einer reellen Form der komplexen* $\boldsymbol{B}_3$.

5. $\nu'(\boldsymbol{B}_4)$ *in einer reellen Form der komplexen* $\boldsymbol{F}_4$.

6. $\nu(\boldsymbol{A}_2)$ *in einer reellen Form der komplexen* $\boldsymbol{G}_2$.

5.8. Wir werden noch zeigen, daß die in Satz IV aufgeführten Einbettungen wirklich vorkommen, und daß sie wesentlich eindeutig bestimmt sind. Wir behandeln erst die elliptischen Fälle.

1. $\boldsymbol{J}=\nu(\boldsymbol{B}_l)$, $\boldsymbol{F}=\boldsymbol{D}_{l+1}$. Man kann einen maximal abelschen Unterring von $\boldsymbol{J}$ mit einem Element von $\boldsymbol{F}$ ergänzen, das zum Gewicht 0 gehört. Das ist dann ein maximal abelscher Unterring von $\boldsymbol{F}$. Man gehe umgekehrt von einem solchen aus. Die Wurzelformen sind $\pm\omega_i\pm\omega_j$. Um $\boldsymbol{J}$ aus $\boldsymbol{F}$ herauszuheben, muß man in dem maximal abelschen Unterring eine lineare Relation einführen, die $4l$ Wurzelformen paarweise zusammenfallen läßt (denn jedes Gewicht $\neq 0$ von $\boldsymbol{J}$ ist einer Wurzelform von $\boldsymbol{J}$ gleich). Im wesentlichen die einzige derartige Relation ist $\omega_{l+1}=0$. (Für $l=3$ kommt noch $\omega_1+\omega_2+\omega_3+\omega_4=0$ hinzu, was aber bekanntlich nichts Neues liefert.) Der maximal abelsche Unterring von $\boldsymbol{J}$ ist hierdurch bestimmt. Um ganz $\boldsymbol{J}$ zu erhalten, muß man noch

$$e_{\pm\omega_i\pm\omega_j},\ \alpha_i e_{\omega_i+\omega_{l+1}}+\beta_i e_{\omega_i-\omega_{l+1}},\ -\bar{\alpha}_i e_{-\omega_i+\omega_{l+1}}-\bar{\beta}_i e_{-\omega_i-\omega_{l+1}} \qquad (i\leq l)$$

hinzufügen. Kommutatorbildung zeigt, daß α_i und β_i im wesentlichen unabhängig von i sein müssen. Durch innere Automorphismen von $\boldsymbol{F}$ können diese $\boldsymbol{J}$ ineinander übergeführt werden; hierzu verwende man $\exp \boldsymbol{h}$ mit einem $\boldsymbol{h}$ aus dem maximal abelschen Unterring, das $\omega_1=\cdots=\omega_l=0$ genügt.

$\boldsymbol{J}=\nu(\boldsymbol{D}_l)$, $\boldsymbol{F}=\boldsymbol{B}_l$. Man muß $2l$ von den Wurzelformen $\pm\omega_i$, $\pm\omega_i\pm\omega_j$ streichen (und als Gewichte erklären), für die paarweise gilt, daß entweder die Summe 0 ist, oder Summe und Differenz eine Wurzelform ist. Das geht nur auf eine Weise.

F ist die Invarianzgruppe $\mathrm{G_0ElR}(n, \mathrm{re})$ von $\sum x_i^2=0$ im projektiven Raum (elliptische Geometrie).

2. $\boldsymbol{J}=\nu(\boldsymbol{A}_l+\boldsymbol{A}_0)$, $\boldsymbol{F}=\boldsymbol{A}_{l+1}$. Der maximal abelsche Unterring von $\boldsymbol{A}_l+\boldsymbol{A}_0$ ist auch einer von $\boldsymbol{A}_{l+1}$. Man gehe von diesem aus. Die Wurzelformen von $\boldsymbol{A}_{l+1}$ seien $\pm\omega_i, \omega_i-\omega_j$. Um $\boldsymbol{J}$ aus $\boldsymbol{F}$ herauszuheben, muß man die Wurzelformen einteilen in Wurzelformen und Gewichte von $\nu(\boldsymbol{A}_l+\boldsymbol{A}_0)$; und zwar müssen zwei Systeme Σ_+, Σ_- von $l+1$ Gewichten auftreten, so daß aus $g\in\Sigma_+$ folgt $-g\in\Sigma_-$ und aus $g, g'\in\Sigma_+$ folgt: $g-g'$ ist Wurzelform. Σ_+ läßt sich im wesentlichen auf zwei Arten wählen:

$$\omega_1, \omega_2, \ldots, \omega_{l+1}$$

oder

$$-\omega_1, \omega_2-\omega_1, \ldots, \omega_{l+1}-\omega_1.$$

Das zweite System entsteht aus dem ersten durch Spiegelung an der Wurzelebene $\omega_1=0$, also durch einen Automorphismus.

F ist die Invarianzgruppe $\mathrm{GElR}(n, \mathrm{ko})$ von $\sum|x_i|^2=0$ im komplexen projektiven Raum (komplex-elliptische Geometrie).

3. $J=\nu(C_l+A_1)$, $F=C_{l+1}$. In bezug auf einen maximal abelschen Unterring von C_l sind die Wurzelformen $\neq 0$ und Gewichte von $\nu(C_l)$:

$$\pm 2\omega_i,\ \pm\omega_i\pm\omega_j \quad \text{und} \quad \pm\omega_i.$$

Die von $\nu(C_l+A_1)$ werden mit dem Parameter ϑ, der von dem maximal abelschen Unterring von A_1 abhängt:

$$\pm 2\omega_i,\ \pm\omega_i\pm\omega_j,\ \pm 2\vartheta \quad \text{und} \quad \pm\omega_i\pm\vartheta.$$

Der maximal abelsche Unterring von C_l+A_1 ist auch einer von C_{l+1}. Wir sondern aus den Wurzelformen von C_{l+1} $4l$ Gewichte ab, die folgende Eigenschaften besitzen müssen: Sind g und g' Gewichte, so ist $g+g'$ oder $g-g'$ Wurzelform; niemals ist $g\pm 2g'$ Wurzelform. Diese Bedingungen bestimmen im wesentlichen eindeutig die Wahl: $\pm\omega_i\pm\omega_{l+1}$.

F ist die Invarianzgruppe GElR(n, qu) von $\sum|x_i|^2=0$ im quaternionprojektiven Raum (quaternion-elliptische Gruppe).

4. $J=\nu(G_2)$, $F=B_3$. Man ergänze einen maximal abelschen Unterring von J mit einem Element von F, das zum Gewicht 0 gehört. Das ist dann ein maximal abelscher Unterring von F. Man gehe von diesem aus. Die Wurzelformen sind $\pm\omega_i$, $\pm\omega_i\pm\omega_j$ $(i, j=1, 2, 3)$. Um J aus F herauszuheben, muß man im maximal abelschen Unterring von F eine lineare Relation einführen, die 12 Wurzelformen paarweise zusammenfallen läßt, aber so, daß keine 0 wird (denn jedes Gewicht $\neq 0$ von J ist einer Wurzelform von J gleich). Im wesentlichen die einzige derartige Relation ist $\omega_1+\omega_2+\omega_3=0$. Der maximal abelsche Unterring von J ist hierdurch bestimmt. Um ganz J zu erhalten, muß man noch

$$e_{\omega_i-\omega_j},\quad \alpha_1 e_{\omega_1+\omega_2}+\beta_1 e_{-\omega_1},\quad \text{usw.}$$

hinzufügen. Man schließt dann wie bei $\nu(B_l)$, D_{l+1} weiter auf die eindeutige Bestimmtheit.

F wird auf der 7-Sphäre induziert durch die Spindarstellung von B_3 im 8-Raum.

5. $J=\nu'(B_4)$, $F=F_4$. Die Wurzelformen $\neq 0$ und Gewichte von $\nu'(B_4)$ sind

$$\pm\omega_i,\ \pm\omega_i\pm\omega_j \quad \text{und} \quad \tfrac{1}{2}(\pm\omega_1\pm\omega_2\pm\omega_3\pm\omega_4).$$

Zusammen bilden sie das Wurzelformensystem der F_4. Geht man umgekehrt von diesem aus, so sieht man, daß die $\pm\omega_i\pm\omega_j$ als Gewichte nicht in Frage kommen, denn zu jedem Gewicht g muß es eine Wurzelform α geben, so daß $\alpha-2g$ Wurzelform ist, und das ist bei $\pm\omega_i\pm\omega_j$ nicht der Fall. Also müssen die $\pm\omega_i\pm\omega_j$ notwendig Gewichte sein. Man darf weiter annehmen, daß auch ω_1 eine Wurzelform ist; dann ergeben sich alle $\pm\omega_i$ als Wurzelformen.

F ist die elliptische Gruppe der Oktavenebene GElR$(2, \mathrm{ok})$.

6. $J=\nu(A_2)$, $F=G_2$. Die Wurzelformen und Gewichte von $\nu(A_2)$ sind

$$\omega_i-\omega_j \quad \text{und} \quad \pm\omega_i \qquad (\omega_1+\omega_2+\omega_3\equiv 0).$$

Das sind auch die Wurzelformen $\neq 0$ von G_2. Zu jeder Wurzelform $\neq 0$ gibt es ein Gewicht g, so daß $\alpha - 3g$ Wurzelform ist. Das ist aber nur bei den $\pm\omega_i \pm \omega_j$ der Fall.

F wird auf der 6-Sphäre induziert von der Darstellung von G_2 als Automorphismengruppe der Oktaven.

Damit ist in allen elliptischen Fällen die Existenz und Eindeutigkeit der Einbettung nachgewiesen.

5.9. Wir untersuchen nun die Einbettungen im hyperbolischen Fall.

$\boldsymbol{F}$ ist nach Satz IV reelle Form einer einfachen Gruppe (von einer Ausnahme abgesehen). Eine reelle Form ist durch ihre Signatur[17]) gekennzeichnet (δ = Anzahl der positiven minus Anzahl der negativen Quadrate der Killing-Cartanschen Form). Die Anzahl der negativen Quadrate ist die maximale Dimension kompakter Untergruppen, also $\geq \dim J$. Die Summe beider Anzahlen ist $\dim F$. Hieraus folgt:

$$\delta \leq \dim \boldsymbol{F} - 2\dim \boldsymbol{J}.$$

Wir suchen in der Cartanschen Aufzählung die reellen Formen, die dieser Ungleichung genügen. Die unitären Formen lassen wir weg.

1. $\boldsymbol{J} = \nu(\boldsymbol{B}_l)$, $\boldsymbol{F} = \boldsymbol{D}_{l+1}$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = (l+1)(2l+1) - 2l(2l+1) = -2l^2 + l + 1.$$

$$\delta = -l - 1;\quad l + 1 - 2m^2 \qquad (m = 0, 1, \ldots, l+1).$$

F ist die Invarianzgruppe $\mathrm{GHyR}(2l+1, \mathrm{re})$ von $\sum_1^{2l+1} x_i^2 - x_{2l+2}^2 = 0$ (Gruppe des hyperbolischen Raumes).

$\boldsymbol{J} = \nu(\boldsymbol{D}_l)$, $\boldsymbol{F} = \boldsymbol{B}_l$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = l(2l+1) - 2l(2l-1) = l(3-2l).$$

$$\delta = l - 2m(m+1) \qquad (m = 0, 1, \ldots, l).$$

F ist die Invarianzgruppe $\mathrm{GHyR}(2l, \mathrm{re})$ von $\sum_1^{2l} x_i^2 - x_{2l+1}^2 = 0$ (Gruppe des hyperbolischen Raumes).

2. $\boldsymbol{J} = \nu(\boldsymbol{A}_l + \boldsymbol{A}_0)$, $\boldsymbol{F} = \boldsymbol{A}_{l+1}$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = (l+3)(l+1) - 2(l+1)^2 = 1 - l^2.$$

$$\delta = l + 1,\ -l - 3[18]);\ 1 - m^2 \qquad (m = l+2, l, \ldots, \geq 0).$$

Für $l = 2$ ist $\delta = -l - 3 = -5$ mit der Ungleichung verträglich. Wir haben dann als reelle Form die Drehgruppe im 6-Raum. Die maximal kompakte Untergruppe ist der B_2 isomorph; sie muß J enthalten. Das kann die

[17]) Bei Cartan „caractère" genannt. Siehe Oeuvres I, 399—491.

[18]) Nur für gerades l. — Man korrigiere bei Cartan, S. 400, Zeile 7:

$\left[\frac{l+3}{2}\right]$ statt $\left[\frac{l+2}{2}\right]$; Zeile 8: $\left[\frac{l+1}{2}\right]$ statt $\left[\frac{l}{2}\right]$.

B_2 aber nicht, denn das würde die Existenz einer treuen Darstellung der B_2 in einer (dim B_2 − dim J)-dimensionalen (= 1-dimensionalen) Mannigfaltigkeit bedeuten. Übrig bleibt:

F ist die Invarianzgruppe GHyR$(l+1, \text{ko})$ von $\sum_1^{l+1} x_i^2 - x_{l+2}^2 = 0$ (Gruppe des komplex hyperbolischen Raumes).

3. $\boldsymbol{J} = \nu(\boldsymbol{C}_l + \boldsymbol{A}_1)$, $\boldsymbol{F} = \boldsymbol{C}_{l+1}$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = (l+1)(2l+3) - 2\big(l(2l+1)+3\big) = -2l^2 + 3l - 3$$

$$\delta = l+1;\quad -l-1-2m^2 \qquad (m = l+1, l-1, \ldots, \geq 0).$$

F ist die Invarianzgruppe GHyR(l, qu) von $\sum_1^{l} x_i^2 - x_{l+1}^2 = 0$ (Gruppe des quaternionhyperbolischen Raumes).

4. $\boldsymbol{J} = \nu(\boldsymbol{G}_2)$, $\boldsymbol{F} = \boldsymbol{B}_3$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = 21 - 28 = -7.$$

$$\delta = 3,\ -1,\ -9,\ -21.$$

Im Falle $\delta = -9$ ist die maximal kompakte Untergruppe der D_3 isomorph, die somit die G_2 enthalten müßte. Das ist unmöglich, denn dann besäße die G_2 eine lineare Darstellung im 6-Raum.

Es gibt hier keine hyperbolische Form.

5. $\boldsymbol{J} = \nu'(\boldsymbol{B}_4)$, $\boldsymbol{F} = \boldsymbol{F}_4$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = 52 - 72 = -20.$$

$$\delta = 4,\ -20,\ -52.$$

F ist die hyperbolische Gruppe der Oktavenebene GHyR(2, ok).

6. $\boldsymbol{J} = \nu(\boldsymbol{A}_2)$, $\boldsymbol{F} = \boldsymbol{G}_2$.

$$\dim \boldsymbol{F} - 2\dim \boldsymbol{J} = 14 - 16 = -2.$$

$$\delta = 2;\ -8,\ 0.$$

Keine hyperbolische Form.

In allen Fällen stimmten die Dimension von J und die der maximal kompakten Untergruppe überein. Hieraus folgt die eindeutige Bestimmtheit der Einbettung.

5.10. Die zulässigen Paare $\boldsymbol{J}, \boldsymbol{F}$ sind hiermit vollständig aufgezählt. Die im elliptischen und hyperbolischen Fall angegebenen Beispiele von Paaren J, F geben eine vollständige Aufzählung im Kleinen bei zusammenhängendem F.

6. Die Verhältnisse im Großen

6.1. Sei F zusammenhängend. Die einfach zusammenhängende Überlagerung heiße $\tilde{F}$. Es gibt eine Untergruppe $\tilde{J}$ von $\tilde{F}$, so daß die kanonische Isomorphie zwischen F/J, $\tilde{F}/\tilde{J}$ und R stattfindet.

Sei $\tilde{J}_0$ die Einskomponente von $\tilde{J}$. Dann ist $\tilde{R} = \tilde{F}/\tilde{J}_0$ die einfach zusammenhängende Überlagerung von R.

Das Paar $\tilde{J}_0, \tilde{F}$ ist durch $\boldsymbol{J}, \boldsymbol{F}$ eindeutig bestimmt. Wir erhalten alle zugehörigen Typen im Großen, wenn wir $\tilde{J}_0$ *innerhalb* $\tilde{F}$ zu einer gemischten Standuntergruppe $\tilde{J}$ mit der Einskomponente $\tilde{J}_0$ erweitern. Das ergibt das Paar $\tilde{J}/\tilde{J} \cap Z, \tilde{F}/\tilde{J} \cap Z$, wenn Z das Zentrum von $\tilde{F}$ ist.

Wir können $\tilde{J}_0, \tilde{F}$ auch ersetzen durch $\tilde{J}_0/\tilde{J}_0 \cap Z, \tilde{F}/\tilde{J}_0 \cap Z$, die wir J^*, F^* nennen, mit dem zugehörigen einfach zusammenhängenden $R^* = F^*/J^*$. Alle zu $\boldsymbol{J}, \boldsymbol{F}$ gehörigen zusammenhängenden Typen ergeben sich dann wieder durch Erweiterung von J^* zu einer gemischten Standuntergruppe mit der Einskomponente J^*. Um zu gegebenem $\boldsymbol{J}, \boldsymbol{F}$ das Paar J^*, F^* zu finden, braucht man nur eine Form im Großen zu wählen mit einfach zusammenhängendem F^*/J^*.

Sei M die Menge der Fixpunkte von J^* in R^*. Sei J eine gemischte Erweiterung von J^* innerhalb F^*. Dann ist J^* Normalteiler von J. Ist $j \in J$ und $m \in M$, so ist $J^* jm = j J^* m = jm$, also $jm \in M$. Also $JM = M$.

Ist J echte Erweiterung von J^*, so muß also M aus mehr als einem Punkt bestehen. J^* vertauscht dann die Punkte von M. Ist J^* „asystatisch", d.h. besitzt J^* nur einen Fixpunkt, so ist keine echte Erweiterung möglich.

6.2. In den parabolischen Fällen nehme man als R^* das abelsche Ideal $\boldsymbol{K}$ von $\boldsymbol{F}$. Als F^* nehme man die von $\exp \boldsymbol{J}$ und der Translationsgruppe des linearen Raumes $\boldsymbol{K}$ erzeugte Gruppe. R^* ist einfach zusammenhängend; $\exp \boldsymbol{J}$ ist die größte kompakte Untergruppe von F^*, also innerhalb F nicht erweiterbar. Zu $\boldsymbol{J}$ gibt es also im Großen nur ein parabolisches Paar J, F.

6.3. In den hyperbolischen Fällen ist R^* ein hyperbolischer Raum, d.h. die Menge

$$\sum_1^n |x_i|^2 < |x_{n+1}|^2$$

im projektiven n-Raum (die x_j sind reell, komplex, Quaternionen oder Oktaven). R^* ist zusammenziehbar:

$$(x_1, \ldots, x_n, x_{n+1}) \to \left(x_1, \ldots, x_n, \frac{1}{1-\lambda} x_{n+1}\right), \qquad 0 \leq \lambda \leq 1.$$

Also ist R^* sicher einfach zusammenhängend. F^* besteht aus den mit der Identität verbindbaren Projektivitäten, die $\sum |x_i|^2 = |x_{n+1}|^2$ invariant lassen; J^* läßt obendrein etwa $(0, 0, \ldots, 1)$ invariant, und das ist der einzige Fixpunkt. Also ist J^* nicht erweiterbar und gibt es im Großen keine anderen Typen.

6.4. Wir behandeln nun die elliptischen Fälle. Der komplex projektive Raum, der quaternion projektive Raum und die projektive Oktavenebene sind einfach zusammenhängend, die Gruppen sind asystatisch. Im Großen kommt also nichts Neues hinzu. Der reelle projektive Raum ist nicht

einfach zusammenhängend; die einfach zusammenhängende Überlagerung ist die der sphärischen Geometrie. J^* hat zwei Fixpunkte, die Erweiterung von J^* führt zur elliptischen Geometrie zurück.

In den übrigen zwei Fällen war R die 6-Sphäre und die 7-Sphäre. J^* besitzt zwei Fixpunkte, die Erweiterung von J^* führt zu entsprechenden Gruppen des projektiven 6-Raumes und 7-Raumes.

7. Der Übergang zu gemischten F

7.1. F_0, die Einskomponente eines F, ist nun vollständig zu übersehen. Nach 2.3 (Anfang des Beweises) erhalten wir F aus F_0, wenn wir $J \cap F_0$ geeignet erweitern, nun aber *außerhalb* F_0. Über die möglichen Erweiterungen von $J \cap F_0$ haben wir ein Bild auf Grund von 4.

7.2. Parabolisch: Die Drehgruppen J werden ergänzt zu orthogonalen Gruppen (mit Spiegelungen). Im unitären Fall kann J eine abgeschlossene Gruppe zwischen $\nu(A_l)$ und $\nu(A_l + A_0)$ sein, noch erweitert mit $x \to \bar{x}$; im quaternion-unitären Fall eine abgeschlossene Gruppe zwischen $\nu(C_l)$ und $\nu(C_l + A_1)$; im oktaven-unitären Fall $(J = \nu'(B_3))$ ist keine Erweiterung möglich. Im Fall $J = \nu(G_2)$ ist Erweiterung mit $x \to -x$ möglich. Im Fall $J = \nu'(B_3)$ gibt es keine Erweiterung.

Alle möglichen Erweiterungen von J sind ohne weiteres zu Erweiterungen von F fortzusetzen.

7.3. Elliptisch: Die sphärischen Geometrien werden mittels einer Spiegelung erweitert. Bei den elliptischen Geometrien gerader Dimension ist das keine echte Erweiterung, bei denen ungerader Dimension wohl. Die unitären Geometrien werden mit $x \to \bar{x}$ erweitert. Für die elliptische Quaternionen- und Oktavengeometrien existiert keine Erweiterung. Von den übrigen besitzt nur die der 6-Sphäre eine Erweiterung.

7.4. Hyperbolisch: Die reelle hyperbolische Geometrie wird mittels einer Spiegelung erweitert, die unitäre mittels $x \to \bar{x}$. Die anderen sind nicht zu erweitern.

7.5. Als Resultat ergibt sich nun die Aufzählung von Hauptsatz I (Einleitung).

8. Ein Nachtrag

8.1. Wir müssen nun noch einen Schluß zu Satz I (s. 2.8) rechtfertigen. Wir zeigen, daß eine Faktorkette von Liegruppen F, wie sie im Hauptsatz I aufgezählt werden, $F, F', F'', \ldots$, in der jede Gruppe echte Faktorgruppe (nach abgeschlossenem Normalteiler) der eventuell vorhergehenden ist, nach höchstens fünf Schritten abbricht.

8.2. Jeder diskrete Normalteiler F_1 eines F besteht aus höchstens zwei Elementen. Denn sei F_0 die Einskomponente von F, $f_1 \in F$. Dann ist die Menge der $f_0 f_1 f_0^{-1}$ mit $f_0 \in F_0$ einerseits zusammenhängend, andererseits diskret,

also einpunktig, also $=(f_1)$. Also f_1 im Zentralisator von F_0. Für jedes $x \in R$ bilde man die Menge $M(x)$ der $x' \in R$ mit der Eigenschaft: aus $fx = x$ $(f \in F_0)$ folgt $fx' = x'$. Dann vertauscht jedes $f_1 \in F_1$ die Elemente von $M(x)$ untereinander. Man sieht in der Aufzählung, daß $M(x)$ für jedes x höchstens zweipunktig ist. Hieraus folgt die Behauptung.

8.3. Wir betrachten in der Faktorkette $F, F', \ldots$ die Dimensionen der auftretenden Gruppen. Ist $\dim F^{(i)} = \dim F^{(i+1)}$, so ist der zugehörige Normalteiler diskret, also von der Mächtigkeit 2; der Kern des Homomorphismus $F^{(i+1)} \to F^{(i+2)}$ kann dann nicht diskret sein. Also $\dim F^{(i+1)} > \dim F^{(i+2)}$. Nun ist ein Absinken der Dimensionen höchstens zweimal möglich; dieses Maximum tritt ein, wenn F oder F' parabolisch und die Faktorgruppe nach dem Abelschen Normalteiler eine Gruppe von der Dimension 6 ist, die selber einen Normalteiler von der Dimension 3 besitzt. Die Länge der Kette ist dann höchstens 5.

Mathematisch Instituut, Rijks-Universiteit, Utrecht (Niederlande)

(Eingegangen am 9. August 1955)

BRAUER, R.
Math. Zeitschr. Bd. 63, S. 406—444 (1956)

Zur Darstellungstheorie der Gruppen endlicher Ordnung

Dem Andenken meines verehrten Lehrers ISSAI SCHUR gewidmet

Von

RICHARD BRAUER

§ 1. Einleitung

In seinen Vorlesungen über Gruppentheorie stellte I. Schur es gelegentlich als wünschenswert dar, die Darstellungen der Gruppen endlicher Ordnung von einem mehr zahlentheoretischen Standpunkt aus zu untersuchen. Die folgenden Ausführungen sind ein Versuch, ein solches Programm in Angriff zu nehmen.

Es sei $\mathfrak{G}$ eine Gruppe endlicher Ordnung g, und $\Gamma = \Gamma(\mathfrak{G}, \Omega)$ sei die zugehörige Gruppenalgebra über einem gegebenen Körper Ω. Das allgemeine Element von Γ hat dann die Form

$$\alpha = \sum_{G \in \mathfrak{G}} a_G G, \tag{1.1}$$

wo die Koeffizienten a_G Elemente aus Ω sind. Wählen wir Ω als algebraischen Zahlkörper endlichen Grades, und ist $\mathfrak{o}$ der Ring der ganzen algebraischen Zahlen in Ω, so bildet die Menge der Elemente (1.1) mit Koeffizienten a_G aus $\mathfrak{o}$ eine Ordnung J in Γ[1]). Man kann dann die Zahlentheorie in J untersuchen. Dabei ist zu beachten, daß J für $g > 1$ nicht Maximalordnung ist, so daß die Ergebnisse von Speiser, Artin und Hasse nicht ohne weiteres angewendet werden können. Zwar ist Γ halbeinfach, und daher kann man J in eine Maximalordnung einbetten, doch gehen dabei gerade die von unserem Standpunkt interessantesten arithmetischen Eigenschaften verloren.

Als erste Frage haben wir die Untersuchung der Primideale $\mathfrak{P}$ von J[2]). Der Durchschnitt von $\mathfrak{P}$ mit $\mathfrak{o}$ ist ein Primideal $\mathfrak{p}$ von $\mathfrak{o}$, und dann ist $J > \mathfrak{P} \supseteqq \mathfrak{p}J$. Die $\mathfrak{p}$ enthaltenden Ideale von J entsprechen eineindeutig den Idealen des Restklassenringes $J^* = J/\mathfrak{p}J$. Die Restklassenabbildung (mod $\mathfrak{p}J$) bildet den Teilring $\mathfrak{o}$ von J auf einen Körper Ω^* ab, der mit $\mathfrak{o}/\mathfrak{p}$ identifiziert werden darf. Man sieht dann, daß man J^* als die Gruppenalgebra $\Gamma(\mathfrak{G}, \Omega^*)$ von $\mathfrak{G}$ über Ω^* auffassen kann;

$$J^* = \Gamma(\mathfrak{G}, \Omega^*); \quad \Omega^* = \mathfrak{o}/\mathfrak{p}. \tag{1.2}$$

---

[1]) Einige Begriffe und Resultate der Algebrentheorie, die im folgenden verwendet werden, sind im Anhang (§ 14) zusammengestellt.

[2]) Vgl. § 14, (a) und (c). Unter einem Ideal wird stets ein *zweiseitiges* Ideal verstanden. Ferner wird von einem Ideal einer Ordnung J vorausgesetzt, daß es von Null verschiedene Elemente von $\mathfrak{o}$ enthält.

Ein $\mathfrak{p}$ enthaltendes Ideal $\mathfrak{P}$ von J ist dann und nur dann Primideal von J, wenn sein Bild $\mathfrak{P}^*$ in J^* Primideal von J^* ist. Aus der Strukturtheorie der Algebren folgt, daß dies dann und nur dann der Fall ist, wenn $\mathfrak{P}^*$ maximal in J^* ist, d.h. wenn $J^*/\mathfrak{P}^*$ eine einfache Algebra ist.

Es seien $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_l$ die wesentlich verschiedenen irreduziblen Darstellungen von $\mathfrak{G}$ durch lineare Transformationen mit Koeffizienten aus Ω^*. Jedes $\mathfrak{F}_i$ kann zu einer Darstellung von $\Gamma(\mathfrak{G}, \Omega^*)$ erweitert werden. Ist $\mathfrak{P}_i^*$ der Kern von $\mathfrak{F}_i$ in diesem Sinn, so sieht man leicht, daß $\mathfrak{P}_1^*, \mathfrak{P}_2^*, \ldots, \mathfrak{P}_l^*$ gerade die verschiedenen Primideale von $\Gamma(\mathfrak{G}, \Omega^*)$ sind. Das $\mathfrak{P}_i^*$ entsprechende Ideal $\mathfrak{P}_i$ von J besteht aus denjenigen Elementen α von J, für deren Restklasse $\alpha^* \pmod{\mathfrak{p}J}$ die Gleichung $\mathfrak{F}_i(\alpha^*) = 0$ gilt. Die so erhaltenen Ideale $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_l$ sind dann die sämtlichen Primidealteiler von $\mathfrak{p}J$ in J. Es ist also die Untersuchung der zu einem festen Primideal $\mathfrak{p}$ von Ω gehörigen Primideale $\mathfrak{P}$ von J äquivalent mit der Untersuchung der irreduziblen Darstellungen von $\mathfrak{G}$ in Ω^*. Mit derartigen „modularen" Darstellungen von endlichen Gruppen haben sich C. J. Nesbitt und der Verfasser in einigen früheren Arbeiten befaßt[3]). In der vorliegenden Arbeit treten die modularen Gruppencharaktere nur als Hilfsmittel auf. Alles, was hier benötigt ist, ist mit kurzen Beweisskizzen in § 3 zusammengestellt.

Will man die Struktur des Ideals $\mathfrak{p}J$ von J weiter verfolgen, so liegt es nahe, die Restklassenalgebra $J^* = \Gamma(\mathfrak{G}, \Omega^*)$ als direkte Summe von direkt unzerlegbaren Idealen $\mathfrak{A}_i^*$ darzustellen;

$$\Gamma(\mathfrak{G}, \Omega^*) = \mathfrak{A}_1^* \oplus \mathfrak{A}_2^* \oplus \cdots \oplus \mathfrak{A}_t^*. \tag{1.3}$$

Setzt man dann

$$\mathfrak{B}_m^* = \sum_{\substack{i=1,2,\ldots,t \\ i \neq m}} \mathfrak{A}_i^*; \quad \mathfrak{A}_m^* = \bigcap_{\substack{i=1,2,\ldots,t \\ i \neq m}} \mathfrak{B}_i^*,$$

und ist $\mathfrak{B}_m$ das $\mathfrak{B}_m^*$ entsprechende Ideal von J, so ist

$$\mathfrak{p}J = \mathfrak{B}_1 \cap \mathfrak{B}_2 \cap \cdots \cap \mathfrak{B}_t. \tag{1.4}$$

Die Ideale $\mathfrak{B}_i$ und $\mathfrak{B}_j$ sind für $i \neq j$ teilerfremd; $(\mathfrak{B}_i, \mathfrak{B}_j) = J$; keins der Ideale $\mathfrak{B}_i$ läßt sich nicht-trivial als Durchschnitt teilerfremder Ideale darstellen. Die Darstellungen (1.3) und (1.4) sind eindeutig bestimmt. Wir nennen $\mathfrak{B}_1, \mathfrak{B}_2, \ldots, \mathfrak{B}_t$ die *Blockideale* von $\mathfrak{p}J$ in J.

Da $\mathfrak{B}_1\mathfrak{B}_2 \ldots \mathfrak{B}_t \subseteq \mathfrak{p}J$ ist, teilt jeder Primidealteiler $\mathfrak{P}$ von $\mathfrak{p}J$ genau eins der Blockideale $\mathfrak{B}_\tau$. Ist also B_τ die Menge der Primidealteiler von $\mathfrak{B}_\tau$, so sind die l Primideale $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_l$ auf t „*Blöcke*" $B_1, B_2, \ldots, B_t$ verteilt. Nach dem oben Gesagten können wir dann auch von einer Verteilung der l modularen Darstellungen $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_l$ auf die t Blöcke sprechen. Wie in § 4 genauer auseinandergesetzt werden wird, führt dies auch zu einer Verteilung der gewöhnlichen irreduziblen Darstellungen von $\mathfrak{G}$ auf die t Blöcke.

[3]) Man vgl. R. Brauer and C. Nesbitt, University of Toronto Studies, Math. Series No. 4, 1937; Annals of Math. (2) 42, 556, Part I.

Diese Blöcke von Darstellungen sollen in der vorliegenden Arbeit untersucht werden[4]).

Kennt man die Charaktere von $\mathfrak{G}$, so kann man daraus weitgehende Information über die Blöcke und die zugehörigen Blockideale ablesen. Dies wird in § 4 und § 5 ausgeführt. Ist p die durch p teilbare rationale Primzahl, und enthält g die Primzahl p genau in der Potenz p^a, so verstehen wir unter dem *Defekt* eines Blocks B die kleinste ganze Zahl d, für die p^{a-d} die Grade aller Darstellungen in B teilt (§ 6).

Das Hauptziel dieser Arbeit ist es, Zusammenhänge zwischen gruppentheoretischen Eigenschaften von $\mathfrak{G}$ und Eigenschaften der Blöcke aufzudecken. Daraus ergeben sich dann neue Eigenschaften der Charaktere von $\mathfrak{G}$. Diese Aufgabe wird in § 7 in Angriff genommen. Auf Grund einer einfachen gruppentheoretischen Bemerkung ergibt sich ein Zusammenhang zwischen den Blöcken gewisser Untergruppen von $\mathfrak{G}$ und den Blöcken von $\mathfrak{G}$. Wir können dann in § 8 jedem Block B von $\mathfrak{G}$ vom Defekt d eine Untergruppe $\mathfrak{D}$ von $\mathfrak{G}$ von der Ordnung p^d, seine *Defektgruppe*, zuordnen. Mit $\mathfrak{D}$ ist auch jede konjugierte Gruppe Defektgruppe von B; davon abgesehen ist $\mathfrak{D}$ eindeutig durch B bestimmt.

Ist $\mathfrak{D}$ eine beliebige Untergruppe von $\mathfrak{G}$, deren Ordnung eine Potenz p^d von p ist und ist $\mathfrak{N}$ der Normalisator von $\mathfrak{D}$ in $\mathfrak{G}$, so gibt es eine eineindeutige Zuordnung zwischen den Blöcken $\tilde{B}$ von $\mathfrak{N}$ vom Defekt d und den Blöcken B von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$ (§ 10). Die Behandlung von $\mathfrak{N}$ läßt sich dann weiter auf Untersuchung von $\mathfrak{N}/\mathfrak{D}$ zurückführen. Für die Werte der Charaktere χ_i in B erhält man Kongruenzen nach Potenzen von p, vgl. (12A).

Interessiert man sich z.B. für die Bestimmung der Anzahl der Blöcke B von einem gegebenen Defekt d, so sieht man jetzt, daß man diese Frage im Fall eines positiven d vollständig auf die Untersuchung von Gruppen von kleinerer Ordnung als g zurückführen kann. Im Fall $d=0$ versagen diese Methoden, da hier $\mathfrak{N}=\mathfrak{G}$, $\mathfrak{D}=\{1\}$ ist. Als Ersatz leiten wir im § 13 ein andersartiges Resultat für die Anzahl der Blöcke vom Defekt 0 her.

Die Untersuchung soll in einer weiteren Mitteilung fortgesetzt werden, während Anwendungen an anderer Stelle behandelt werden sollen.

§ 2. Bezeichnungen

Im folgenden ist $\mathfrak{G}$ stets eine endliche Gruppe der Ordnung g. Ist $\mathfrak{A}$ eine beliebige Teilmenge von $\mathfrak{G}$, so bezeichnen wir mit $\mathfrak{N}(\mathfrak{A})$ den *Normalisator* von $\mathfrak{A}$, d.h. die Untergruppe von $\mathfrak{G}$, deren Elemente X die Gleichung $X\mathfrak{A}=\mathfrak{A}X$ erfüllen. Ferner sei $\mathfrak{C}(\mathfrak{A})$ der *Zentralisator* von $\mathfrak{A}$, das ist die Untergruppe von $\mathfrak{G}$, deren Elemente mit jedem Element von $\mathfrak{A}$ vertauschbar sind. Besteht $\mathfrak{A}$ aus einem einzigen Element A, so schreiben wir $\mathfrak{N}(A)$ $(=\mathfrak{C}(A))$ für $\mathfrak{N}(\mathfrak{A})$. Die Ordnung von $\mathfrak{N}(A)$ werde mit $n(A)$ bezeichnet.

[4]) Über einen Teil der Ergebnisse berichten ohne Beweis drei Noten, Proc. Nat. Acad. Sci. **30**, 109—114 (1944); **32**, 182—186, 215—219 (1946).

Es seien im folgenden $\mathfrak{K}_1, \mathfrak{K}_2, \ldots, \mathfrak{K}_k$ stets die Klassen konjugierter Elemente von $\mathfrak{G}$. Ist also $G_\alpha \in \mathfrak{K}_\alpha$, so besteht $\mathfrak{K}_\alpha$ aus $g/n(G_\alpha)$ Elementen. Die Anzahl k der Klassen $\mathfrak{K}_\alpha$ ist dann auch die Anzahl der irreduziblen Charaktere von $\mathfrak{G}$, die mit $\chi_1, \chi_2, \ldots, \chi_k$ bezeichnet werden sollen. Es sei x_i der Grad von χ_i; $x_i = \chi_i(1)$.

Ist Ξ ein beliebiger Körper, so kann man die Gruppenalgebra $\Gamma(\mathfrak{G}, \Xi)$ von $\mathfrak{G}$ über Ξ bilden. Dies ist eine assoziative Algebra der Dimension g über Ξ, deren g Basiselemente so gewählt werden können, daß sie bei Multiplikation eine zu $\mathfrak{G}$ isomorphe Gruppe bilden. Man kann dann die Basiselemente mit den Elementen von $\mathfrak{G}$ identifizieren, und dann ist das allgemeine Element von $\Gamma(\mathfrak{G}, \Xi)$ von der Form (1.1) mit Koeffizienten $a_G \in \Xi$. Offenbar ist das Einheitselement von $\mathfrak{G}$ auch Einheitselement von $\Gamma(\mathfrak{G}, \Xi)$. Mit K_α werde stets das Element von $\Gamma(\mathfrak{G}, \Xi)$ bezeichnet, das die Summe der Elemente der Klasse $\mathfrak{K}_\alpha$ ist,

$$K_\alpha = \sum_{G \in \mathfrak{K}_\alpha} G. \tag{2.1}$$

Bekanntlich bilden die k Elemente $K_1, K_2, \ldots, K_k$ eine Basis des Zentrums $Z(\mathfrak{G}, \Xi)$ der Gruppenalgebra. Wegen dieses Zusammenhangs nennen wir $Z(\mathfrak{G}, \Xi)$ die *Klassenalgebra* von $\mathfrak{G}$ über Ξ. Die Multiplikation der Basiselemente ist durch Formeln

$$K_\alpha K_\beta = \sum_\gamma a_{\alpha\beta\gamma} K_\gamma \tag{2.2}$$

gegeben. Hier ist $a_{\alpha\beta\gamma}$ eine nicht-negative ganze rationale Zahl, die von der Wahl des Körpers unabhängig ist. Ist G_γ ein Element von $\mathfrak{K}_\gamma$, so ist nämlich $a_{\alpha\beta\gamma}$ die Anzahl der geordneten Paare (X, Y) von $\mathfrak{G}$, die die Bedingungen

$$XY = G_\gamma, \quad X \in \mathfrak{K}_\alpha, \quad Y \in \mathfrak{K}_\beta$$

erfüllen. Hat Ξ Primzahlcharakteristik p, so kann man natürlich $a_{\alpha\beta\gamma} \pmod p$ nehmen.

Ist Ξ der Körper der komplexen Zahlen, und ist $\mathfrak{X}_i$ die zum irreduziblen Charakter χ_i von $\mathfrak{G}$ gehörige Darstellung von $\mathfrak{G}$, so kann man $\mathfrak{X}_i$ als irreduzible Darstellung von $\Gamma(\mathfrak{G}, \Xi)$ auffassen. Ist ζ ein beliebiges Element von $Z(\mathfrak{G}, \Xi)$, so ist $\mathfrak{X}_i(\zeta)$ nach dem Schurschen Lemma ein skalares Vielfaches $\omega_i(\zeta) E$ der Identität E; $\omega_i(\zeta) \in \Xi$. Als Funktion von ζ stellt dann ω_i einen linearen Charakter der Klassenalgebra $Z(\mathfrak{G}, \Xi)$ dar. Natürlich ist ω_i vollständig durch die Werte $\omega_i(K_\alpha)$ für $\alpha = 1, 2, \ldots, k$ bestimmt. Diese kann man in bekannter Weise durch Berechnung der Spur von $\mathfrak{X}_i(K_\alpha)$ bestimmen. Ist G_α ein Element von $\mathfrak{K}_\alpha$, so wird es oft bequem sein, auch $\omega_i(G_\alpha)$ für $\omega_i(K_\alpha)$ zu schreiben; man hat dann

$$\omega_i(K_\alpha) = \omega_i(G_\alpha) = \frac{g}{n(G_\alpha)} \frac{\chi_i(G_\alpha)}{x_i}, \quad \left(x_i = \chi_i(1)\right). \tag{2.3}$$

Da $\chi_1, \chi_2, \ldots, \chi_k$ orthogonal und daher linear unabhängig sind, sind die k linearen Charaktere $\omega_1, \omega_2, \ldots, \omega_k$ von $Z(\mathfrak{G}, \Xi)$ voneinander verschieden. Da

$Z(\mathfrak{G}, \Xi)$ die Dimension k hat, sind dies sämtliche irreduziblen Darstellungen der Klassenalgebra. Man sieht auch jetzt noch leicht ein, daß $Z(\mathfrak{G}, \Xi)$ halbeinfach ist.

Jedes $\chi_i(G_\alpha)$ ist eine Summe von g-ten Einheitswurzeln. Ersetzt man also den Körper der komplexen Zahlen Ξ durch einen die g-ten Einheitswurzeln enthaltenden Teilkörper Ω, so zeigt (2.3), daß alle $\omega_i(K_\alpha)$ in Ω liegen. Daher besitzt auch $Z(\mathfrak{G}, \Omega)$ genau k verschiedene lineare Charaktere, die wir ebenfalls mit $\omega_1, \omega_2, \ldots, \omega_k$ bezeichnen.

Im folgenden ist Ω stets ein die g-ten Einheitswurzeln enthaltender algebraischer Zahlkörper endlichen Grades. Wie in § 1 sei $\mathfrak{o}$ der Ring der ganzen algebraischen Zahlen in Ω, und es sei $\mathfrak{p}$ ein festes Primideal von $\mathfrak{o}$ und p die durch $\mathfrak{p}$ teilbare rationale Primzahl. Die Restklasse eines Elements α von $\mathfrak{o}$ (mod $\mathfrak{p}$) wird stets mit α^* bezeichnet werden; es sei

$$\Omega^* = \mathfrak{o}/\mathfrak{p} \tag{2.4}$$

der Restklassenkörper. Ferner sei ν stets die zu $\mathfrak{p}$ gehörige exponentielle Bewertung von Ω, die wir so normieren, daß

$$\nu(p) = 1 \tag{2.5}$$

ist. Gelegentlich wird die Bezeichnung $\mathfrak{o}_\nu$ für den Ring der für $\mathfrak{p}$ ganzen Zahlen von Ω und $\mathfrak{p}_\nu$ für das Primideal von $\mathfrak{o}_\nu$ verwendet werden. Der Restklassenkörper $\mathfrak{o}_\nu/\mathfrak{p}_\nu$ darf mit Ω^* identifiziert werden.

Ist G ein Element von $\mathfrak{G}$, so schreiben wir zur Abkürzung $\nu(G)$ für $\nu\big(n(G)\big)$ und nennen diese Zahl den *Defekt* von G. Die p-Sylowgruppen des Normalisators $\mathfrak{N}(G)$ haben dann also die Ordnung $p^{\nu(G)}$. Gehört G_α zur Klasse $\mathfrak{K}_\alpha$, so setzen wir noch

$$\nu(\mathfrak{K}_\alpha) = \nu(G_\alpha) = \nu\big(n(G_\alpha)\big), \qquad (G_\alpha \in \mathfrak{K}_\alpha); \tag{2.6}$$

wir nennen $\nu(\mathfrak{K}_\alpha)$ den *Defekt* der Klasse $\mathfrak{K}_\alpha$.

Ein Element G von $\mathfrak{G}$ heiße *p-regulär*, wenn die Ordnung von G zu p teilerfremd ist. Da das gleiche dann auch für alle konjugierten Elemente gilt, so können wir von *p-regulären Klassen* konjugierter Elemente sprechen.

§ 3. Eigenschaften von modularen Charakteren

Wir geben in diesem Abschnitt eine Schilderung der Theorie der modularen Gruppendarstellungen, soweit sie im folgenden benötigt wird.

(3A) *Es sei Γ eine Algebra von endlicher Dimension über einem algebraisch abgeschlossenen Körper Ξ von Primzahlcharakteristik p. Es sei S der von den Elementen $\alpha\beta - \beta\alpha$ mit $\alpha, \beta \in \Gamma$ aufgespannte lineare Teilraum von Γ; es sei T die Menge der Elemente ϱ von Γ, für die es eine Potenz q von p gibt, derart daß ϱ^q zu S gehört. Dann ist T ein linearer Teilraum von Γ. Ferner ist die Anzahl der wesentlich verschiedenen irreduziblen Darstellungen von Γ gleich der Dimension des Restklassenraumes Γ/T.*

Beweis. Es seien ϱ und σ zwei Elemente von Γ; man bilde $(\varrho+\sigma)^p$ durch Ausmultiplizieren der Klammern. Zu jedem von ϱ^p und σ^p verschiedenen Glied gehören p Glieder, die aus ihm durch zyklische Permutation der Faktoren entstehen. Offenbar sind diese p Glieder untereinander kongruent (mod S). Daher ergibt sich

$$(\varrho+\sigma)^p \equiv \varrho^p + \sigma^p \pmod{S}. \tag{3.1}$$

Sind α, β Elemente von Γ und setzt man $\gamma = \beta(\alpha\beta)^{p-1}$, so folgt unter Verwendung von (3.1), daß

$$(\alpha\beta - \beta\alpha)^p \equiv (\alpha\beta)^p - (\beta\alpha)^p = \alpha\gamma - \gamma\alpha \equiv 0 \pmod{S}$$

ist. Jetzt sieht man, daß die p-te Potenz jedes Elementes von S wieder zu S gehört. Unter Verwendung von (3.1) folgt nun leicht, daß die in (3A) definierte Menge T wirklich ein linearer Teilraum von Γ ist, und ferner daß $T \supseteq S$ ist.

Ist Γ einfach, also eine volle Matrizenalgebra über Ξ, so besteht S, wie man leicht sieht, gerade aus den Matrizen von der Spur Null. Ist η ein Idempotent mit von Null verschiedener Spur, so gehört η nicht zu S und auch nicht zu T. Da nun Γ/S die Dimension 1 hat und $\Gamma \supset T \supseteq S$ ist, so folgt hier, daß Γ/T die Dimension 1 hat.

Im allgemeinen Fall folgt aus der Nilpotenz des Radikals N von Γ, daß N zu T gehört. Besitzt nun Γ genau l wesentlich verschiedene irreduzible Darstellungen, so ist Γ/N direkte Summe von l Idealen Γ_i, die einfache Algebren sind. Hat T_i die entsprechende Bedeutung für Γ_i wie T für Γ, so folgt jetzt leicht, daß die Dimension von Γ/T gleich der Summe der Dimensionen der Γ_i/T_i ist. Nach dem oben Gesagten ist aber jede der letzteren Dimensionen 1, und daher hat Γ/T die Dimension l wie behauptet war.

(3B) *Die Anzahl der wesentlich verschiedenen irreduziblen Darstellungen einer endlichen Gruppe $\mathfrak{G}$ in einem algebraisch abgeschlossenen Körper Ξ von Primzahlcharakteristik p ist gleich der Anzahl der p-regulären Klassen konjugierter Elemente in $\mathfrak{G}$.*

Beweis. Wir nehmen $\Gamma = \Gamma(\mathfrak{G}, \Xi)$ und verwenden dieselben Bezeichnungen wie im Beweis von (3A). Ist G ein beliebiges Element von $\mathfrak{G}$, so können wir $G = RP$ setzen, wo R und P Potenzen von G sind und wo R p-regulär ist, während die Ordnung von P eine Potenz q von p ist. Aus (3.1) mit $\varrho = RP$, $\sigma = -R$ ergibt sich

$$(RP - R)^p \equiv R^p P^p - R^p \pmod{S}.$$

Diese Beziehung erhebe man wieder in die p-te Potenz und beachte, daß wie oben bemerkt, die p-te Potenz eines Elementes von S wieder zu S gehört. Fährt man so fort, so findet man schließlich

$$(RP - R)^q \equiv R^q P^q - R^q = R^q - R^q = 0 \pmod{S}.$$

Dies heißt aber, daß $RP - R \in T$ ist, also $G \equiv R \pmod{T}$.

Jetzt folgt, daß (mod T) das allgemeine Element (1.1) von $\Gamma(\mathfrak{G}, \Xi)$ einem Element derselben Form kongruent ist, in dem nur die p-regulären Elemente von $\mathfrak{G}$ von Null verschiedene Koeffizienten haben. Sind ferner G und G_1 zwei konjugierte Elemente von $\mathfrak{G}$, so ist $G \equiv G_1 \pmod{S}$ und also auch (mod T). Daher kann man setzen

$$\alpha \equiv \sum_R a_R R \pmod{T}, \tag{3.2}$$

wo R ein Vertretersystem für die p-regulären Klassen durchläuft. Diese Elemente R bilden ein (mod T) linear unabhängiges System von Elementen von Γ. Ist nämlich α in (3.2) ein Element von T, so gibt es eine Potenz q von p, derart daß $\alpha^q \in S$ ist. Da die Elemente R p-regulär sind, kann man q so wählen, daß für jedes R die Gleichung $R^q = R$ gilt. Erhebt man nun α in (3.2) in die q-te Potenz, so ergibt sich mit Hilfe von wiederholter Anwendung von (3.1), daß

$$\sum_R a_R^q R \equiv 0 \pmod{S}$$

ist. Nun sieht man aber leicht, daß S aus den Elementen (1.1) besteht, für die (mod p) für jede Klasse $\mathfrak{K}_i$ die Summe der Koeffizienten der Elemente G in $\mathfrak{K}_i$ verschwindet. Da alle R in (3.2) zu verschiedenen Klassen gehören, sind (mod p) alle a_R^q und daher alle a_R in (3.2) Null, wie zu zeigen war.

Es bilden also die Elemente R in (3.2) eine Basis für $\Gamma \pmod{T}$. Daher ist die Dimension l von Γ/T gleich der Anzahl der p-regulären Klassen $\mathfrak{K}_i$, und nun liefert (3A) die Behauptung.

Wir beweisen weiter den für uns wichtigen Hilfssatz

(3C) *Sind $\mathfrak{K}_1, \mathfrak{K}_2, \ldots, \mathfrak{K}_l$ die l p-regulären Klassen von $\mathfrak{G}$, und ist G_j ein Repräsentant für $\mathfrak{K}_j$, so kann man l (gewöhnliche) irreduzible Charaktere $\chi_1, \chi_2, \ldots, \chi_l$ von $\mathfrak{G}$ so auswählen, daß*

$$\operatorname{Det}\left(\chi_i(G_j)\right) \not\equiv 0 \pmod{\mathfrak{p}} \tag{3.3}$$

ist, wobei in der Determinante der Zeilenindex i und der Spaltenindex j die Werte $1, 2, \ldots, l$ durchlaufen.

Beweis. Wäre (3C) nicht richtig, so könnte man l ganze Zahlen $u_1, u_2, \ldots, u_l$ in Ω so wählen, daß für jeden irreduziblen Charakter χ_i die Kongruenz gilt

$$\sum_{j=1}^{l} u_j \chi_i(G_j) \equiv 0 \pmod{\mathfrak{p}}$$

und etwa $u_m \not\equiv 0 \pmod{\mathfrak{p}}$ ist. Ist Θ eine beliebige Linearverbindung der χ_i mit ganzen Koeffizienten in Ω, so ist dann

$$\sum_{j=1}^{l} u_j \Theta(G_j) \equiv 0 \pmod{\mathfrak{p}}. \tag{3.4}$$

Es sei nun $\mathfrak{B} = \{G_m\}$ die von G_m erzeugte zyklische Gruppe, b sei die Ordnung und $\zeta_1, \zeta_2, \ldots, \zeta_b$ seien die b irreduziblen Charaktere von $\mathfrak{B}$. Dann wähle

man eine p-Sylowgruppe $\mathfrak{Q}$ von $\mathfrak{N}(G_m)$ und setze $\mathfrak{E}=\{\mathfrak{B},\mathfrak{Q}\}=\mathfrak{B}\times\mathfrak{Q}$. Ist BQ mit $B\in\mathfrak{B}$, $Q\in\mathfrak{Q}$ ein variables Element von $\mathfrak{E}$, so ist ein Charakter ϑ_i von $\mathfrak{E}$ durch

$$\vartheta_i(BQ)=\zeta_i(B)$$

definiert. Es sei Θ_i der von ϑ_i induzierte Charakter von $\mathfrak{G}$. Setzt man

$$\Theta=\sum_{i=1}^{b}\overline{\zeta_i(G_m)}\,\Theta_i,$$

so ist Θ eine Linearverbindung von Charakteren von $\mathfrak{G}$ mit ganzen Koeffizienten aus Ω. Aus der Definition der induzierten Charaktere ergibt sich

$$\Theta(G_j)=\frac{1}{(\mathfrak{E}:1)}\sum_{X\in\mathfrak{G}}\sum_{i=1}^{b}\overline{\zeta_i(G_m)}\,\vartheta_i(XG_jX^{-1}),$$

wo $\vartheta_i(XG_jX^{-1})$ gleich Null zu setzen ist, falls XG_jX^{-1} nicht zu $\mathfrak{E}$ gehört. Ist $XG_jX^{-1}\in\mathfrak{E}$ und $1\leq j\leq l$, so ist XG_jX^{-1} p-regulär und also ein Element B von $\mathfrak{B}$. Dann ist $\vartheta_i(XG_jX^{-1})=\zeta_i(B)$. Aus den Orthogonalitätsrelationen für die Charaktere von $\mathfrak{B}$ ergibt sich jetzt, daß der Wert der inneren Summe in der letzten Formel 0 oder b ist, je nachdem ob $B\neq G_m$ oder $B=G_m$ ist. Das letztere ist nur für $G_j=G_m$, d.h. für $j=m$ möglich. Jetzt folgt leicht, daß $\Theta(G_j)=0$ für $j\neq m$ und $\Theta(G_m)=b(\mathfrak{N}(G_m):\mathfrak{E})$ ist. Wegen der Art, in der wir $\mathfrak{E}$ konstruierten, ist dieser Wert $\Theta(G_m)$ eine zu p teilerfremde ganze Zahl. Setzen wir dies in (3.4) ein, so erhalten wir einen Widerspruch, da $u_m\not\equiv 0 \pmod{\mathfrak{p}}$ war. Damit ist (3C) bewiesen.

Da der Körper Ω nach Voraussetzung die g-ten Einheitswurzeln enthält, kann man die irreduzible Darstellung $\mathfrak{X}_i$ von $\mathfrak{G}$ mit dem Charakter χ_i in Ω schreiben, $i=1,2,\ldots,k$[5]). Ebenso kann man die Koeffizienten der l absolut irreduziblen modularen Darstellungen $\mathfrak{F}_1,\mathfrak{F}_2,\ldots,\mathfrak{F}_l$ von $\mathfrak{G}$ im Restklassenkörper Ω^* wählen. Ist $\mathfrak{o}_\nu$ der Ring der für $\mathfrak{p}$ ganzen Zahlen in Ω, so darf man ferner annehmen, daß $\mathfrak{X}_i(G)$ für jedes $G\in\mathfrak{G}$ Koeffizienten in $\mathfrak{o}_\nu$ hat[6]). Die Restklassenabbildung führt dann $\mathfrak{X}_i$ in eine Darstellung $\mathfrak{X}_i^*$ mit Koeffizienten in $\mathfrak{o}_\nu/\mathfrak{p}_\nu=\Omega^*$ über. Tritt die irreduzible modulare Darstellung $\mathfrak{F}_j$ in $\mathfrak{X}_i$ mit der Vielfachheit d_{ij} auf, so nennen wir die nicht-negativen ganzen rationalen Zahlen d_{ij} die *Zerlegungszahlen* von $\mathfrak{G}$; $i=1,2,\ldots,k$; $j=1,2,\ldots,l$.

Wir definieren jetzt den Charakter φ einer modularen Darstellung $\mathfrak{F}$ mit Koeffizienten in Ω^*. Dabei beschränken wir uns auf p-reguläre Elemente G von $\mathfrak{G}$. Hat G die Ordnung h, so sind die charakteristischen Wurzeln $\varepsilon_1^*,\varepsilon_2^*,\ldots$ von $\mathfrak{F}(G)$ sämtlich h-te Einheitswurzeln. Da h zu p teilerfremd ist, geht

[5]) Vgl. etwa R. Brauer, Amer. J. Math. **69**, 709 (1947). Ersetzt man gegebenfalls Ω durch einen Erweiterungskörper endlichen Grades, so kann man die Verwendung dieser Tatsache vermeiden.

[6]) Vgl. § 14, (d) und (14B). Dort wird auch gezeigt, daß die im folgenden definierten Zahlen d_{ij} ungeändert bleiben, wenn man $\mathfrak{X}_i$ durch eine ähnliche Darstellung mit Koeffizienten in $\mathfrak{o}_\nu$ ersetzt.

jedes ε_i^* aus einer eindeutig bestimmten h-ten Einheitswurzel ε_i in Ω durch Restklassenabbildung hervor. Dann setzen wir $\varphi(G) = \sum \varepsilon_i$. Es ist also φ eine für p-reguläre Elemente $G \in \mathfrak{G}$ definierte Funktion mit Werten in Ω. Natürlich hängt $\varphi(G)$ nur von der Klasse konjugierter Elemente $\mathfrak{K}_\alpha$ ab, zu der G gehört.

(3D) *Haben die irreduziblen modularen Darstellungen $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_l$ die Charaktere $\varphi_1, \varphi_2, \ldots, \varphi_l$, sind d_{ij} die Zerlegungszahlen, so ist für p-reguläre Elemente G*

$$\chi_i(G) = \sum_{j=1}^{l} d_{ij} \varphi_j(G). \tag{3.5}$$

In der Tat stimmt für p-reguläres G der Wert des Charakters von $\mathfrak{X}_i^*$ mit dem Wert $\chi_i(G)$ des Charakters χ_i von $\mathfrak{X}_i$ überein, und da $\mathfrak{F}_j$ genau d_{ij}-mal in $\mathfrak{X}_i^*$ auftritt, so erhält man sofort (3.5).

Setzt man die Werte (3.5) in (3.3) ein, so erhält man

(3E) *Die aus den Werten der l irreduziblen modularen Charaktere φ_i für die Repräsentanten $G_1, G_2, \ldots, G_l$ der l p-regulären Klassen von $\mathfrak{G}$ gebildete Determinante ist nicht durch $\mathfrak{p}$ teilbar;*

$$\operatorname{Det}\left(\varphi_i(G_j)\right) \not\equiv 0 \pmod{\mathfrak{p}}. \tag{3.6}$$

Sind ferner $\chi_1, \chi_2, \ldots, \chi_l$ wie in (3C) gewählt, so ist

$$\operatorname{Det}(d_{ij}) \not\equiv 0 \pmod{p}; \qquad (i, j = 1, 2, \ldots, l). \tag{3.7}$$

Wir setzen noch

$$c_{ij} = \sum_{\varrho=1}^{h} d_{\varrho i} d_{\varrho j} = c_{ji}, \qquad (i, j = 1, 2, \ldots, l); \tag{3.8}$$

$$\Phi_i = \sum_{j=1}^{h} d_{ji} \chi_j, \qquad (i = 1, 2, \ldots, l). \tag{3.9}$$

Die c_{ij} sind nicht-negative ganze rationale Zahlen, die sog. *Cartanschen Invarianten* von $\mathfrak{G}$ (oder eigentlich von $\Gamma(\mathfrak{G}, \Omega^*)$). Es erübrigt sich hier, ihre algebrentheoretische Bedeutung zu diskutieren. Die Φ_i sind gewöhnliche Charaktere von $\mathfrak{G}$, die in der modularen Theorie eine besondere Rolle spielen, worauf wir ebenfalls nicht näher einzugehen brauchen. Aus (3.9) und (3.5) ergibt sich unter Verwendung von (3.8) die Beziehung

$$\Phi_i(G) = \sum_{j=1}^{l} c_{ij} \varphi_j(G) \qquad (G \ p\text{-regulär}). \tag{3.10}$$

Die Orthogonalitätsbeziehungen für die Charaktere von $\mathfrak{G}$ besagen, daß

$$\sum_{i=1}^{h} \overline{\chi_i(G)}\, \chi_i(H) = n(G) \quad \text{oder} \quad = 0,$$

je nachdem ob die Elemente G und H von $\mathfrak{G}$ konjugiert sind oder nicht. Wählt man für H eins der p-regulären Elemente $G_1, G_2, \ldots, G_l$, sagen wir

$H = G_\alpha$, und drückt man $\chi_i(G_\alpha)$ nach (3.5) durch $\varphi_j(G_\alpha)$ aus, so folgt mit Hilfe von (3.9), daß

$$\sum_{j=1}^{l} \overline{\Phi_j(G)}\, \varphi_j(G_\alpha) = n(G) \quad \text{oder} \quad = 0 \tag{3.11}$$

ist, je nachdem ob G und G_α konjugiert sind oder nicht. Für p-singuläres G tritt der zweite Fall ein. Aus dem Nichtverschwinden der Determinante (3.6) ergibt sich dann, daß $\Phi_j(G) = 0$ sein muß. Für p-reguläres G folgt ebenso aus (3.6), daß $\overline{\Phi_j(G)}$ durch die höchste Potenz von $\mathfrak{p}$ teilbar ist, die in $n(G)$ aufgeht. Wegen $\overline{\Phi_j(G)} = \Phi_j(G^{-1})$, $n(G) = n(G^{-1})$ haben wir

(3 F) *Die in* (3.9) *definierten Charaktere Φ_j verschwinden für alle p-singulären Elemente. Für p-reguläres G ist $\nu\big(\Phi_j(G)\big) \geq \nu(G)$.*

Ist G p-regulär, so kann man $\Phi_j(G)$ in (3.11) nach (3.10) durch $\varphi_i(G)$ ausdrücken. Schreibt man für G_α wieder H, so erhält man

(3 G) *Für p-reguläre Elemente G und H von $\mathfrak{G}$ gilt*

$$\sum_{j=1}^{l} \overline{\Phi_j(G)}\, \varphi_j(H) = \sum_{i,j=1}^{l} c_{ij}\, \overline{\varphi_i(G)}\, \varphi_j(H) = n(G) \quad \text{oder} \quad 0, \tag{3.12}$$

je nachdem ob G und H konjugiert sind oder nicht.

§ 4. Blöcke von Charakteren

Wir schicken unserer Untersuchung einige Bemerkungen über die Klassenalgebra $Z = Z(\mathfrak{G}, \Omega)$ voraus. Wie bereits in § 2 bemerkt wurde, besitzt Z die k verschiedenen linearen Charaktere $\omega_1, \omega_2, \ldots, \omega_k$ und da Z die Dimension k hat, sind dies die sämtlichen irreduziblen Darstellungen von Z.

Offenbar ist der Durchschnitt $J \cap Z = J_0$ eine Ordnung von Z, die die $\mathfrak{o}$-Basis $K_1, K_2, \ldots, K_k$ besitzt; der Restklassenring $J_0/\mathfrak{p} J_0$ darf mit $Z^* = Z(\mathfrak{G}, \Omega^*)$ identifiziert werden. Anwendung der Restklassenabbildung auf $\omega_1, \omega_2, \ldots, \omega_k$ liefert dann k lineare Charaktere $\omega_1^*, \omega_2^*, \ldots, \omega_k^*$ von Z^*, die aber im allgemeinen nicht alle verschieden sind. Jeder Homomorphismus ψ der Algebra Z^* über Ω^* auf einen Erweiterungskörper von Ω^* ist ein linearer Charakter von Z^* und stimmt mit einem der ω_i^* überein[7]).

Wie in (1.3) betrachten wir die Zerlegung der modularen Gruppenalgebra $\Gamma(\mathfrak{G}, \Omega^*)$ in direkt unzerlegbare Ideale

$$\Gamma(\mathfrak{G}, \Omega^*) = \mathfrak{A}_1^* \oplus \mathfrak{A}_2^* \oplus \cdots \oplus \mathfrak{A}_t^*. \tag{4.1}$$

Setzt man dementsprechend

$$1 = \eta_1^* + \eta_2^* + \cdots + \eta_t^*, \qquad (\eta_\tau^* \in \mathfrak{A}_\tau^*), \tag{4.2}$$

so sind die η_τ^* orthogonale Idempotente des Zentrums $Z^* = Z(\mathfrak{G}, \Omega^*)$ von $\Gamma(\mathfrak{G}, \Omega^*)$. Das Ideal $\eta_\tau^* Z^*$ von Z^* ist direkt unzerlegbar und man hat

$$Z^* = \eta_1^* Z^* \oplus \eta_2^* Z^* \oplus \cdots \oplus \eta_t^* Z^*.$$

[7]) Vgl. § 14; (14 A) und (14 C).

Als kommutative direkt unzerlegbare Algebra ist $\eta_\tau^* Z^*$ primär. Ist also $\mathfrak{B}_\tau^*$ das Radikal von $\eta_\tau^* Z^*$, so ist $\eta_\tau^* Z^*/\mathfrak{B}_\tau$ ein Körper, der als Erweiterungskörper Ω_τ^* von Ω^* betrachtet werden kann. Man erweitere den natürlichen Homomorphismus von $\eta_\tau^* Z^*$ auf Ω_τ^* zu einem Homomorphismus ψ_τ von Z^* auf Ω_τ^*, indem man $\psi_\tau(\xi)=0$ für $\xi \in \eta_\sigma^* Z^*$ mit $\sigma \neq \tau$ vorschreibt. Nach dem oben Gesagten erhält man so einen linearen Charakter ψ_τ von Z^*; jeder lineare Charakter von Z^* stimmt mit einem dieser ψ_τ überein. Man hat $\psi_\tau(\eta_\tau^*)=1$.

Wie in § 1 sei $\mathfrak{B}_\tau^*$ die Summe der $\mathfrak{A}_\sigma^*$ mit $\sigma \neq \tau$. Da die Ideale $\mathfrak{B}_\tau^*$ von $\Gamma(\mathfrak{G}, \Omega^*)$ die Bilder der Ideale $\mathfrak{B}_\tau$ von $\Gamma(\mathfrak{G}, \Omega)$ in (1.4) sind, so haben wir das Resultat

(4A) *Hat das Ideal $\mathfrak{p}J$ in J die Blockideale $\mathfrak{B}_1, \mathfrak{B}_2, \ldots, \mathfrak{B}_t$ und bildet der natürliche Homomorphismus von J auf $J/\mathfrak{p}J = \Gamma(\mathfrak{G}, \Omega^*)$ das Ideal $\mathfrak{B}_\tau$ auf das Ideal $\mathfrak{B}_\tau^*$ von $\Gamma(\mathfrak{G}, \Omega^*)$ ab, so gibt es lineare Charaktere ψ_τ von $Z(\mathfrak{G}, \Omega^*)$, derart daß $\psi_\tau(\eta_\tau^*)=1$ ist; $\tau = 1, 2, \ldots, t$. Die linearen Charaktere $\psi_1, \psi_2, \ldots, \psi_t$ sind die sämtlichen linearen Charaktere der Algebra $Z(\mathfrak{G}, \Omega^*)$.*

Ist wie immer $\mathfrak{p}$ ein festes Primideal von Ω, so ordnen wir jedem Blockideal $\mathfrak{B}_\tau$ von $\mathfrak{p}J$ einen „*Block*" B_τ zu. Darunter werden wir eine gewisse Menge von gewöhnlichen und modularen irreduziblen Darstellungen von $\mathfrak{G}$ verstehen. An Stelle von Darstellungen werden wir auch von den zu B_τ gehörigen gewöhnlichen und modularen Charakteren von $\mathfrak{G}$ sprechen. Den linearen Charakter ψ_τ von Z in (4A) nennen wir *den dem Block B_τ zugeordneten linearen Charakter der modularen Klassenalgebra $Z^* = Z(\mathfrak{G}, \Omega^*)$.*

Wir rechnen eine gewöhnliche irreduzible Darstellung $\mathfrak{X}_i$ von $\mathfrak{G}$ und ihren Charakter χ_i zum Block B_τ, wenn für den χ_i nach (2.3) entsprechenden linearen Charakter ω_i von Z die Gleichung gilt

$$\omega_i^* = \psi_\tau. \tag{4.3}$$

Aus dem oben Gesagten ergibt sich

(4B) *Jeder irreduzible Charakter χ_i von $\mathfrak{G}$ gehört zu einem und nur einem Block B_τ und jeder Block enthält Charaktere χ_i. Zwei Charaktere χ_i und χ_j gehören dann und nur dann zum selben Block, wenn für alle $G \in \mathfrak{G}$ die Kongruenz gilt*

$$\omega_i(G) \equiv \omega_j(G) \pmod{\mathfrak{p}}. \tag{4.4}$$

Wir zählen eine modulare irreduzible Darstellung $\mathfrak{F}_j$ und ihren Charakter φ_j zum Block B_τ, wenn der Kern $\mathfrak{P}_j^*$ von $\mathfrak{F}_j$ das Ideal $\mathfrak{B}_\tau^*$ enthält.

(4C) *Jede irreduzible modulare Darstellung $\mathfrak{F}_j$ gehört zu einem und nur einem Block B_τ und jeder Block B_τ enthält derartige $\mathfrak{F}_j$. Gehört die gewöhnliche irreduzible Darstellung $\mathfrak{X}_i$ zu B_τ, so gehören alle in $\mathfrak{X}_i^*$ als Bestandteil auftretenden $\mathfrak{F}_j$ zu B_τ.*

Beweis. Der erste Teil von (4C) ergibt sich unmittelbar aus dem in § 1 Gesagten. Gehört $\mathfrak{X}_i$ zu B_τ, so gilt (4.3). Wegen $\psi_\tau(\eta_\tau^*)=1$ ist $\omega_i^*(\eta_\tau^*)=1$. Als Element von Z^* hat η_τ^* die Form $\sum b_\lambda^* K_\lambda$ mit $b_\lambda^* \in \Omega^*$. Wählt man b_λ

in $\mathfrak{o}$ derart daß die Restklassenabbildung von $\mathfrak{o}$ auf Ω^* gerade b_λ in b_λ^* überführt und setzt man $\eta_\tau = \sum b_\lambda K_\lambda$, so ist η_τ ein Element von $J_0 = Z \cap J$, das durch die Restklassenabbildung von J gerade in η_τ^* übergeht. Dann ist $\mathfrak{X}_i(\eta_\tau) = \omega_i(\eta_\tau) E$, also $\mathfrak{X}_i^*(\eta_\tau^*) = \omega_i^*(\eta_\tau^*) E^* = E^*$. Tritt nun $\mathfrak{F}_j$ als Bestandteil in $\mathfrak{X}_i^*$ auf, so ist $\mathfrak{F}_j(\eta_\tau^*) = E^*$. Also ist $\eta_\tau^* \notin \mathfrak{P}_j^*$ und daher ist $\mathfrak{B}_\sigma$ für $\sigma \neq \tau$ nicht in $\mathfrak{P}_j$ enthalten. Dann muß $\mathfrak{B}_\tau \subseteq \mathfrak{P}_j$ sein, und daher gehört $\mathfrak{F}_j$ zum Block B_τ, wie zu zeigen war.

Aus (4C) ersieht man, daß die Zerlegungszahl d_{ij} verschwinden muß, wenn $\mathfrak{X}_i$ und $\mathfrak{F}_j$ zu verschiedenen Blöcken gehören. Daraus folgt unter Verwendung von (3.8), daß $c_{ij} = 0$ ist, wenn $\mathfrak{F}_i$ und $\mathfrak{F}_j$ zu verschiedenen Blöcken gehören. Dies Resultat läßt sich folgendermaßen aussprechen

(4D) *Bildet man aus den Zerlegungszahlen d_{ij} und den Cartanschen Invarianten c_{ij} Matrizen D und C, so sind bei geeigneter Anordnung der Indizes diese Matrizen direkte Summen von den den einzelnen Blöcken B_τ entsprechenden Teilmatrizen D_τ bzw. C_τ.*

Besteht B_τ aus k_τ gewöhnlichen Darstellungen $\mathfrak{X}_i$ und aus l_τ modularen Darstellungen $\mathfrak{F}_j$, so hat D_τ gerade k_τ Zeilen und l_τ Spalten. Die Matrix C_τ ist quadratisch vom Grade l_τ. Aus (3.8) ergibt sich

$$C_\tau = D_\tau' D_\tau, \tag{4.5}$$

wo D_τ' die Transponierte von D_τ ist.

Beschränkt man sich ausschließlich auf die Betrachtung der gewöhnlichen Darstellungen, so hängt die Verteilung auf Blöcke allein von der rationalen Primzahl p und nicht von dem Primidealteiler $\mathfrak{p}$ von p ab. Beim Beweis dieser Behauptung darf man annehmen, daß Ω normal über dem Körper Q der rationalen Zahlen ist. Jeder Primidealteiler $\mathfrak{p}'$ von p in Ω geht aus dem festen Primideal $\mathfrak{p}$ durch Anwendung eines Elements T der Galoisschen Gruppe von Ω über Q hervor. Ist nun ε eine primitive g-te Einheitswurzel, so führt T dann ε in eine Potenz ε^s mit einem zu g teilerfremden Exponenten s über. Da $\chi_i(G)$ eine Summe von g-ten Einheitswurzeln ist, wird $\chi_i(G)$ in $\chi_i(G^s)$ transformiert, also $\omega_i(G)$ in $\omega_i(G^s)$. Wendet man T auf (4.4) an und ersetzt G^s durch G, so sieht man, daß (4.4) für $\mathfrak{p}'$ an Stelle von $\mathfrak{p}$ gilt. Daraus ergibt sich die Behauptung.

Kennt man die gewöhnlichen irreduziblen Charaktere χ_i von $\mathfrak{G}$, so kann man für jedes Primideal sofort die Anzahl t der Blöcke B_τ bestimmen, d.h. also feststellen, wieviel Blockideale $\mathfrak{B}_\tau$ in der Zerlegung (1.4) von $\mathfrak{p}J$ auftreten. Die Anzahl der in $\mathfrak{B}_\tau$ aufgehenden Primideale $\mathfrak{P}_j$ von J ist die Anzahl l_τ der modularen Darstellungen $\mathfrak{F}_j$ in B_τ. Sind die modularen Charaktere φ_j bekannt, so kann man leicht l_τ bestimmen. Die unten beim Beweis von (5A) verwendete Methode zeigt, daß man l_τ auch finden kann wenn man nur die χ_i kennt. Da unter unseren Voraussetzungen der Restklassenring $J/\mathfrak{P}_j$ ein voller Matrizenring über Ω^* ist, ist seine Struktur durch seine Gradzahl f_j bestimmt. Es ist aber f_j der Grad der zu $\mathfrak{P}_j$ gehörigen modularen Darstellung $\mathfrak{F}_j$, also $f_j = \varphi_j(1)$. Da sich die Cartanschen Invarianten c_{ij} aus

den modularen Charakteren berechnen lassen[8]), so erhält man auch gewisse Information bezüglich des Restklassenrings $J/\mathfrak{B}_\tau$. Kennt man die modularen Charaktere φ_j, so kann man auch ein System von Kongruenzen für die Koeffizienten a_G des Elements α in (1.1) aufstellen, die die notwendigen und hinreichenden Bedingungen darstellen, daß α zu dem Primideal $\mathfrak{P}_j$ gehört. Es sei $\varphi_i^*(G)$ die Spur von $\mathfrak{F}_i(G)$ für $G \in \mathfrak{G}$. Setzt man wie gelegentlich in § 3 $G = RP$, wo R und P Potenzen von G sind, und wo R p-regulär ist, während die Ordnung von P eine Potenz von p ist, so sieht man leicht, daß $\varphi_i^*(G) = \varphi_i^*(R)$ ist, und daß $\varphi_i^*(R)$ aus dem Werte $\varphi_i(R)$ des Charakters φ_i von $\mathfrak{F}_i$ für das p-reguläre Element R durch Restklassenabbildung hervorgeht. Dann gilt

(4E) *Das zur irreduziblen modularen Darstellung $\mathfrak{F}_i$ von $\mathfrak{G}$ gehörige Primideal $\mathfrak{P}_i$ von J besteht aus denjenigen Elementen $\alpha = \sum a_G G$ von J, $a_G \in \mathfrak{o}$, für die die Gleichungen*

$$\sum_G a_G^* \varphi_i^*(GH) = 0 \tag{4.6}$$

für alle $H \in \mathfrak{G}$ gelten. Hier bezeichnet a_G^* das aus $a_G \in \mathfrak{o}$ durch Restklassenabbildung (mod $\mathfrak{p}$) *erhaltene Element von $\Omega^* = \mathfrak{o}/\mathfrak{p}$.*

Beweis. Man sieht leicht, daß die Elemente $\alpha \in J$, für die Gln. (4.6) gelten, ein Ideal $\mathfrak{M}$ von J bilden; $\mathfrak{P}_i \subseteq \mathfrak{M} \subseteq J$. Daher ist entweder $J = \mathfrak{M}$ oder $J = \mathfrak{P}_i$. Wäre $J = \mathfrak{M}$, so müßte die Spur von $\mathfrak{F}_i(\alpha^*)$ für alle $\alpha^* \in J^*$ verschwinden, was nicht der Fall ist. Daher ist $\mathfrak{M} = \mathfrak{P}_i$, wie behauptet war.

Der Beweis gilt auch dann noch, wenn Ω ein beliebiger algebraischer Zahlkörper ist, also nicht notwendig die g-ten Einheitswurzeln enthält. In der Tat ist dann Ω^* ein endlicher Körper, also vollkommen, und daher enthält J^* Elemente α^*, denen in der irreduziblen Darstellung $\mathfrak{F}_i$ von J^* lineare Transformationen $\mathfrak{F}_i(\alpha^*)$ mit von Null verschiedener Spur entsprechen.

§ 5. Zuordnung von p-regulären Klassen zu den Blöcken

Enthält der Block B_τ wie zuvor l_τ modulare Charaktere und ist l die Anzahl der p-regulären Klassen von $\mathfrak{G}$, so ist nach (3B)

$$l = l_1 + l_2 + \cdots + l_t. \tag{5.1}$$

Wir wollen jetzt die l p-regulären Klassen so den Blöcken B_τ zuordnen, daß dem Block B_τ gerade l_τ Klassen entsprechen, $\tau = 1, 2, \ldots, t$. Dabei wählen wir die Bezeichnungen so, daß dieselben Indizes j bei den $\varphi_j \in B_\tau$ und bei den B_τ zugeordneten Klassen auftreten. Genauer wollen wir zeigen

---

[8]) Sieht man die Zahlen $n(G)$ für p-reguläres G als bekannt an, so sieht man dies aus (3.12) und (3.6). Deutlicher noch ergibt sich diese Behauptung aus den Formeln $\gamma_{ij} g = \sum_R \overline{\varphi_i(R)}\, \varphi_j(R)$, wo R über alle p-regulären Elemente läuft und wo (γ_{ij}) die zu $C = (c_{ij})$ inverse Matrix C^{-1} bezeichnet. Für den Beweis der Formeln sei auf die in [3]) zitierten Arbeiten verwiesen.

(5A) *Es sei S_τ die Menge der l_τ Indizes j, für die φ_j zu B_τ gehört, $\tau = 1, 2, \ldots, t$. Man kann die gewöhnlichen Charaktere χ_i von $\mathfrak{G}$ und die Klassen $\mathfrak{K}_i$ so numerieren, daß alle Klassen $\mathfrak{K}_i$ mit $i \in S_\tau^*$ p-regulär sind, daß die Charaktere χ_i mit $i \in S_\tau$ zu B_τ gehören und daß für jedes τ die folgenden Beziehungen gelten*

$$\left.\begin{aligned} &\text{(5.2)} && \operatorname{Det}(\chi_i(G_j)) \not\equiv 0 \pmod{\mathfrak{p}}, \\ &\text{(5.3)} && \operatorname{Det}(\varphi_i(G_j)) \not\equiv 0 \pmod{\mathfrak{p}}, \\ &\text{(5.4)} && \operatorname{Det}(d_{ij}) \not\equiv 0 \pmod{p}, \end{aligned}\right\} \quad (i, j \in S_\tau),$$

wobei in allen drei Determinanten der Zeilenindex i und der Spaltenindex j über die l_τ Werte in S_τ laufen.

Beweis. Nach (3C) kann man l Charaktere χ_i so auswählen, daß

$$\text{(5.5)} \qquad \Delta = \operatorname{Det}(\chi_i(G_j)) \not\equiv 0 \pmod{\mathfrak{p}}$$

ist, wobei hier i und j die Werte $1, 2, \ldots, l$ durchlaufen und $G_1, G_2, \ldots, G_l$ Vertreter für die l p-regulären Klassen $\mathfrak{K}_1, \mathfrak{K}_2, \ldots, \mathfrak{K}_l$ sind. Entwickelt man Δ nach dem Laplaceschen Satz, so kann man Δ in der Form schreiben

$$\Delta = \sum_{\varrho} \pm \Delta_1^{(\varrho)} \Delta_2^{(\varrho)} \ldots \Delta_t^{(\varrho)},$$

wobei $\Delta_\tau^{(\varrho)}$ eine Unterdeterminante ist, in der die zu B_τ gehörigen unter den Charakteren χ_i auftreten. Es gebe etwa m_τ derartige Charaktere. Ist $m_\tau = 0$, so ist $\Delta_\tau^{(\varrho)} = 1$ zu setzen. Da in Δ gerade l Charaktere χ_i auftreten, hat man

$$\text{(5.6)} \qquad l = m_1 + m_2 + \cdots + m_t.$$

Wegen (5.5) kann man ein Glied $\Delta_1^{(\varrho)} \Delta_2^{(\varrho)} \ldots \Delta_t^{(\varrho)}$ so auswählen, daß

$$\text{(5.7)} \qquad \Delta_1^{(\varrho)} \Delta_2^{(\varrho)} \ldots \Delta_t^{(\varrho)} \not\equiv 0 \pmod{\mathfrak{p}}$$

ist. Drückt man die in $\Delta_\tau^{(\varrho)}$ auftretenden Charaktere nach (3.5) durch modulare Charaktere φ_j aus, so treten nur modulare Charaktere $\varphi_j \in B_\tau$ auf, vgl. (4D). Wäre nun der Grad m_τ von $\Delta_\tau^{(\varrho)}$ größer als die Anzahl l_τ der modularen Charaktere in B_τ, so würde $\Delta_\tau^{(\varrho)}$ verschwinden, was im Widerspruch zu (5.7) steht. Daher ist $m_\tau \leq l_\tau$ für jedes τ. Vergleich von (5.1) und (5.6) zeigt jetzt, daß $m_\tau = l_\tau$ ist, $\tau = 1, 2, \ldots, t$.

Ändert man jetzt die Numerierung der Charaktere χ_i und der Klassen $\mathfrak{K}_j$ so ab, daß für jedes τ in der Determinante $\Delta_\tau^{(\varrho)}$ gerade die l_τ Charaktere χ_i mit $i \in S_\tau$ und die l_τ Klassen $\mathfrak{K}_j$ mit $j \in S_\tau$ auftreten, so folgt (5.2) aus (5.7). Drückt man wie oben die χ_i in dieser Determinante durch die φ_i aus, so sieht man, daß die Determinante in (5.2) das Produkt der Determinanten in (5.3) und (5.4) ist. Daher sind auch diese beiden Determinanten nicht durch $\mathfrak{p}$ teilbar. Da die d_{ij} ganze rationale Zahlen sind, kann man in (5.4) den Modul $\mathfrak{p}$ durch p ersetzen, und damit ist (5A) bewiesen. Die Bedingungen von (5A) können im allgemeinen auf mehrere Weisen erfüllt werden.

Nach (3.12) ist

$$\sum_{i,j} c_{\alpha\beta}\,\overline{\varphi_i(G_\alpha)}\,\varphi_j(G_\beta) = n(G_\alpha)\,\delta_{\alpha\beta},$$

wobei $\delta_{\alpha\beta}=0$ oder 1 zu setzen ist, je nachdem ob $\alpha \neq \beta$ oder $\alpha=\beta$ ist; $\alpha, \beta = 1, 2, \ldots, l$. Ist φ die Matrix $(\varphi_i(G_j))$, C die Matrix (c_{ij}) und N die Matrix $(n(G_i)\delta_{ij})$, $(i, j = 1, 2, \ldots, l)$, so kann man die Gleichung folgendermaßen in Matrizenform schreiben

$$\overline{\varphi}' C \varphi = N. \tag{5.8}$$

Arbeitet man im Ring $\mathfrak{o}_\mathfrak{p}$ der für $\mathfrak{p}$ ganzen Zahlen von Ω, so ist die Determinante von φ nach (3.6) eine Einheit. Daher zeigt (5.8), daß die Elementarteiler von C die in $n(G_1), n(G_2), \ldots, n(G_l)$ auftretenden Potenzen einer lokalen Primzahl von $\mathfrak{o}_\mathfrak{p}$ sind. Statt dessen kann man dann auch die Potenzen

$$p^{\nu(G_1)}, p^{\nu(G_2)}, \ldots, p^{\nu(G_l)} \tag{5.9}$$

nehmen. Da C ganze rationale Koeffizienten hat, sind dies auch die Elementarteiler von C im Ring $I_\mathfrak{p}$ der für $\mathfrak{p}$ ganzen, rationalen Zahlen. Tatsächlich ist die Determinante von C eine Potenz von p[9]), und daher sind die Zahlen (5.9) die Elementarteiler von C im Ring der ganzen rationalen Zahlen. Es wird aber im folgenden genügen, im Ring $I_\mathfrak{p}$ zu operieren. Wir zeigen

(5B) *Die Elementarteiler der zum Block B_τ gehörigen Teilmatrix C_τ von C sind die Zahlen $p^{\nu(G_j)}$ mit $j \in S_\tau$.*

Beweis. Hat C_τ die Elementarteiler $e_1, e_2, \ldots, e_{l_\tau}$, so zeigen wir zunächst, daß bei geeigneter Anordnung der j in S_τ die Ungleichungen $e_j \geq p^{\nu(G_j)}$ gelten. Wir können die e_j als Potenzen von p wählen und setzen dementsprechend $e_j = p^{a_j}$. Ist H_τ eine Diagonalmatrix, die in der Hauptdiagonale die Elementarteiler e_j enthält, so gibt es zwei in $I_\mathfrak{p}$ unimodulare Matrizen $U=(u_{ij})$ und $V=(v_{ij})$, $(i, j \in S_\tau)$, vom Grade l_τ derart, daß

$$U C_\tau V^{-1} = H_\tau \tag{5.10}$$

ist.

Nach (3.10) und (4D) haben wir für p-reguläres G

$$\Phi_i(G) = \sum_{j \in S_\tau} c_{ij}\,\varphi_j(G); \qquad (i \in S_\tau). \tag{5.11}$$

Man setze

$$\tilde{\Phi}_i(G) = \sum_{j \in S_\tau} u_{ij}\,\Phi_j(G); \quad \tilde{\varphi}_i(G) = \sum_{j \in S_\tau} v_{ij}\,\varphi_j(G); \qquad (i \in S_\tau). \tag{5.12}$$

Wegen (5.10) nimmt dann (5.11) die Form an

$$\tilde{\Phi}_j(G) = e_j\,\tilde{\varphi}_j(G) = p^{a_j}\tilde{\varphi}_j(G), \qquad (j \in S_\tau).$$

---

[9]) Vgl. R. BRAUER, Ann. of Math. 42, 53—61 (1941) oder Ann. of Math. 57, 357—377 (1953), Satz 13.

Jetzt zeigt (3F), daß für alle p-regulären G und für alle $j \in S_\tau$ die Ungleichungen gelten

$$\nu\left(\tilde{\varphi}_j(G)\right) \geq \nu(G) - \varepsilon_j. \tag{5.13}$$

Aus (5.12) erhält man

$$\mathrm{Det}\left(\tilde{\varphi}_i(G_j)\right) = \mathrm{Det}\,(v_{ij})\,\mathrm{Det}\left(\varphi_i(G_j)\right); \qquad (i, j \in S_\tau).$$

Da V unimodular ist, ist $\mathrm{Det}\,(v_{ij})$ eine Einheit von I_p. Aus (5.3) folgt daher, daß die Determinante auf der linken Seite eine Einheit von $\mathfrak{o}_\tau$ ist. Nimmt man also die l_τ Elemente G_j mit $j \in S_\tau$ in geeigneter Anordnung, so darf man annehmen, daß

$$\nu\left(\prod_{j \in S_\tau} \tilde{\varphi}_j(G_j)\right) = 0$$

ist. Dann ist $\nu\left(\tilde{\varphi}_j(G_j)\right) = 0$, und jetzt zeigt (5.13), daß $\varepsilon_j \geq \nu(G_j)$ ist, wie wir es zeigen wollten. Diese Ungleichungen gelten für jeden Block B_τ. Da C die direkte Summe der Matrizen C_τ ist, ist die Menge der Elementarteiler von C die Vereinigungsmenge der Mengen der Elementarteiler der C_τ. Würde auch nur in einem einzigen Fall in $\varepsilon_j \geq \nu(G_j)$ das Ungleichheitszeichen gelten, so könnten wir nicht in (5.9) die Elementarteiler von C haben. Also gilt stets $\varepsilon_j = \nu(G_j)$ und damit ist (5B) vollständig bewiesen.

Unsere Methode liefert auch noch den Hilfssatz

(5C) *Ist p^m der größte Elementarteiler von C_τ, so gelten für alle $\varphi_i \in B_\tau$, alle $\chi_i \in B_\tau$ und alle p-regulären Elemente G von $\mathfrak{G}$ die Beziehungen*

$$\nu\left(\varphi_i(G)\right) \geq \nu(G) - m; \qquad \nu\left(\chi_i(G)\right) \geq \nu(G) - m. \tag{5.14}$$

In der Tat folgt aus (5.13), daß $\nu\left(\tilde{\varphi}_j(G)\right) \geq \nu(G) - m$ ist, $j \in S_\tau$. Da V unimodular war, kann man nach den zweiten Gln. (5.12) φ_i als Linearverbindung der $\tilde{\varphi}_j$ mit Koeffizienten in $\mathfrak{o}_\tau$ darstellen. Daraus folgt die erste Behauptung (5.14). Die zweite Behauptung folgt aus der ersten unter Beachtung von (3.5) und (4D).

§ 6. Der Defekt eines Blockes

Es sei jetzt p^a die höchste in g aufgehende Potenz der Primzahl p, also in unseren Bezeichnungen $a = \nu(g)$. Wir definieren den Defekt d eines Blockes B_τ als die kleinste ganze rationale Zahl d, für die die Grade x_i aller Charaktere $\chi_i \in B_\tau$ durch p^{a-d} teilbar sind. Man kann dann also $x_i = p^{a-d} x_i'$ mit ganzem rationalem x_i' setzen; es gibt $\chi_i \in B_\tau$, für die x_i' zu p teilerfremd ist.

Wir haben bei der Definition des Defekts die gewöhnlichen Charaktere von B_τ verwendet. Wir hätten ebensogut die modularen Charaktere nehmen können. In der Tat gilt

(6A) *Hat B_τ den Defekt d, so sind die Grade f_i aller modularen irreduziblen Charaktere φ_i in B_τ sämtlich durch p^{a-d}, aber nicht alle durch p^{a-d+1} teilbar.*

Beweis. Man bestimme die ganze rationale Zahl δ so, daß die Grade f_i der $\varphi_i \in B_\tau$ alle durch $p^{a-\delta}$, aber nicht alle durch $p^{a-\delta+1}$ teilbar sind. Aus

(3.5) ergibt sich für $G=1$ unter Berücksichtigung von (4D), daß $p^{a-\delta}$ ein Teiler von p^{a-d} ist. Andererseits folgt aus (5.4) und (3.5), daß für p-reguläre Elemente die $\varphi_i \in B_\tau$ als Linearverbindungen von l_τ geeigneten Charakteren χ_i aus B_τ mit rationalen, für p ganzen, Koeffizienten darstellen kann. Für $G=1$ zeigt dies, daß p^{a-d} ein Teiler von $p^{a-\delta}$ ist. Daher muß $d=\delta$ sein, wie behauptet war.

Es seien jetzt Determinanten $\Delta_1^{(\varphi)}, \Delta_2^{(\varphi)}, \ldots, \Delta_l^{(\varphi)}$ wie in (5.7) ausgewählt. Verwenden wir dieselben Bezeichnungen wie in § 5, so haben wir also

$$\Delta_\tau^{(\varphi)} = \mathrm{Det}\left(\chi_i(G_j)\right) \not\equiv 0 \pmod{\mathfrak{p}}, \tag{6.1}$$

wo i und j die l_τ Werte in S_τ durchlaufen. Man wähle eine feste Spalte, etwa die r-te Spalte und multipliziere die Elemente dieser Spalte mit $g/n(G_r)$. Nach (2.3) ist $g\,\chi_i(G_r)/n(G_r) = x_i\,\omega_i(G_r)$. Da x_i durch p^{a-d} teilbar ist und $\omega_i(G_r)$ ganz algebraisch ist, so ist $g\,\Delta_\tau^{(\varphi)}/n(G_r)$ durch p^{a-d} teilbar. Aus (6.1) folgt dann

$$\nu(G_r) \leq d. \tag{6.2}$$

Da die χ_i mit $i \in S_\tau$ alle zum Block B_τ gehören, so sind nach (4.4) die l_τ Zahlen $\omega_i(G_r)$ für die l_τ Werte $i \in S_\tau$ alle untereinander (mod $\mathfrak{p}$) kongruent. Daher kann das Gleichheitszeichen in (6.2) höchstens dann gelten, wenn $\omega_i(G_r) \not\equiv 0 \pmod{\mathfrak{p}}$ für $i \in S_\tau$ ist.

Nehmen wir nun an, daß für alle $r \in S_\tau$ das Ungleichheitszeichen in (6.2) gilt. Aus (5B) folgt dann, daß alle Elementarteiler von C kleiner als p^d sind. Dann ist $m \leq d-1$ in (5.14) und für $G=1$ ergibt sich, daß für die Grade aller $\chi_i \in B_\tau$ die Ungleichung $\nu(x_i) > a-d$ gilt. Dies widerspricht der Definition des Defekts d. Wir können also $r \in S_\tau$ so wählen, daß

$$\nu(G_r) = d, \qquad \omega_i(G_r) \not\equiv 0 \pmod{\mathfrak{p}} \tag{6.3}$$

ist.

Wählt man, was nach Definition von d möglich ist, den Charakter $\chi_i \in B_\tau$ zunächst so, daß $\nu(x_i) = a-d$ ist, so hat man

$$\nu\left(\omega_i(G)\right) = \nu\left(\frac{g\,\chi_i(G)}{n(G)\,x_i}\right) = a + \nu\left(\chi_i(G)\right) - \nu(G) - \nu(x_i) \geq d - \nu(G).$$

Insbesondere ist $\omega_i(G) \equiv 0 \pmod{\mathfrak{p}}$, falls $\nu(G) < d$ ist. Wegen (4.4) gilt dies für alle $i \in S_\tau$. Sagen wir wie in § 2, daß ein Element G den Defekt ν_0 hat, wenn $\nu(G) = \nu_0$ ist, so liefert (6.3) das folgende Resultat

(6B) *Ein irreduzibler Charakter χ_i von $\mathfrak{G}$ gehört dann und nur dann zu einem Block vom Defekt d, wenn für alle Elemente G von $\mathfrak{G}$ von einem Defekt kleiner als d die Kongruenz $\omega_i(G) \equiv 0 \pmod{\mathfrak{p}}$ gilt, während es p-reguläre Elemente G vom Defekt d gibt, für die $\omega_i(G) \not\equiv 0 \pmod{\mathfrak{p}}$ ist. (Wie immer ist ω_i der zu χ_i gehörige Charakter der Klassenalgebra $Z(\mathfrak{G}, \Omega)$.)*

Es sei jetzt B_τ ein Block, für den $l_\tau \geq 2$ ist und es seien r und s zwei verschiedene zu S_τ gehörige Indizes. Multipliziert man dann in (6.1) die r-te Spalte mit $g/n(G_r)$, die s-te Spalte mit $g/n(G_s)$ und drückt man in beiden

Spalten wieder die χ_i nach (2.3) durch die ω_i aus, so erhält man eine Formel

$$\frac{g}{n(G_r)}\,\frac{g}{n(G_s)}\,\Delta_\tau = \left|\ldots,\ x_i\omega_i(G_r),\ldots,x_i\omega_i(G_s),\ldots\right|,$$

wo auf der rechten Seite die i-te Zeile der Determinante angegeben ist, und wo in den nicht hingeschriebenen Spalten ganze algebraische Koeffizienten, nämlich Werte $\chi_i(G_j)$ stehen. Da alle $\chi_i \in B_\tau$ Grade x_i haben, die durch p^{a-d} teilbar sind, so kann man aus der r-ten und der s-ten Spalte je den Faktor p^{a-d} herausziehen. Ist dies geschehen, und wendet man auf die so entstehende Determinante die Restklassenabbildung (mod $\mathfrak{p}$) an, so werden nach (4.4) zwei Spalten proportional; man erhält also 0. Nach (6.1) hat man also $a-\nu(G_r)+a-\nu(G_s)>2(a-d)$, d.h.

$$\nu(G_r)+\nu(G_s)<2d. \tag{6.4}$$

Ist G_r wie in (6.3) gewählt, so hat man $\nu(G_s)<d$. Es ist also genau einer der Werte $\nu(G_j)$ mit $j\in S_\tau$ gleich d und alle anderen sind kleiner als d. Auf Grund von (5B) liefert dies

(6C) *Ist C_τ die zu einem Block vom Defekt d gehörige Teilmatrix der Matrix der Cartanschen Invarianten von $\mathfrak{G}$, so hat C_τ genau einen Elementarteiler p^d, während alle Elementarteiler kleiner als p^d sind.*

Da nach (5.5) die Elementarteiler aller C_τ zusammen bekannt sind, liefert dies Ergebnis bereits gewisse Auskunft über die Anzahl der Blöcke von gegebenem Defekt[10]). Insbesondere kann eine Zahl p^a in (5.9) nur von einem Block vom Defekt a herrühren, und jeder solche Block liefert nur einmal p^a. Wir haben also

(6D) *Die Anzahl der Blöcke B vom maximalen Defekt $a=\nu(g)$ ist gleich der Anzahl der p-regulären Klassen $\mathfrak{K}_j$ vom Defekt a.*

Im Fall $d=0$ kann C_τ überhaupt nur einen einzigen Elementarteiler haben, und dieser muß 1 sein. Man hat also $l_\tau=1$, $C_\tau=(1)$. Nach (4.5) folgt dann, daß D_τ auch nur eine einzige Zeile haben kann, da alle d_{ij} ganze rationale Zahlen sind, und keine Zeile von D_τ ausschließlich aus Nullen besteht, wie man z.B. aus (3.5) für $G=1$ sieht. So erhalten wir das Resultat

(6E) *Jeder Block vom Defekt 0 besteht aus genau einem gewöhnlichen irreduziblen Charakter und einem irreduziblen modularen Charakter. Diese beiden Charaktere stimmen für p-reguläre Elemente überein, während der erstere für alle nicht p-regulären Elemente verschwindet.*

Der letzte Teil ergibt sich aus (3F) unter Verwendung von (3.9) und $D_\tau=(1)$.

Gehört χ_i zu einem Block B vom Defekt 0, so nennen wir χ_i einen *Charakter vom Defekt* 0. Nach Definition ist der Grad x_i durch die höchst mögliche

[10]) Die folgenden Ergebnisse (6D) und (6E) finden sich in der zweiten in [3]. erwähnten Arbeit. Der Satz (6C) wurde von C. Nesbitt und dem Verfasser kurz nach dem Entstehen der Arbeiten [3]. mit einem von dem hier gegebenen verschiedenen Beweis gefunden aber nicht veröffentlicht.

Potenz p^a von p teilbar. Man kann auch die Umkehrung zeigen, daß aus $p^a | x_i$ folgt, daß χ_i zu einem Block vom Defekt 0 gehört[11]). Für eine derartige Darstellung $\mathfrak{X}_i$ ist die entsprechende modulare Darstellung $\mathfrak{X}_i^*$ irreduzibel und nach den Formeln (3.9) gleichzeitig ein direkter Summand der modularen regulären Darstellung von $\mathfrak{G}$. Daraus folgt nach der allgemeinen Theorie der Darstellungen von Algebren auch noch, daß $\mathfrak{X}_i^*$ ein direkter Summand in allen modularen Darstellungen ist, in denen es als Bestandteil auftritt. Wir werden die letzteren Tatsachen im folgenden nicht verwenden. Man sieht, daß im Fall des Defekts 0 die modulare Darstellungstheorie mit der gewöhnlichen Darstellungstheorie zusammenfällt. Im allgemeinen sind die Abweichungen um so größer, je größer der Defekt ist. Es sei noch bemerkt, daß im Fall $g \not\equiv 0 \pmod{p}$ alle Blöcke natürlich den Defekt 0 haben müssen und daß daher überhaupt kein Unterschied zwischen der modularen und der gewöhnlichen Theorie besteht.

Von Wichtigkeit für uns ist das folgende Kriterium

(6F) *Es seien χ_α und χ_β zwei irreduzible Charaktere von $\mathfrak{G}$, die beide zu Blöcken vom Defekt d gehören. Die beiden Charaktere gehören dann und nur dann zu demselben Block, wenn für alle p-regulären Elemente G vom Defekt d die Kongruenz gilt*

$$\omega_\alpha(G) \equiv \omega_\beta(G) \pmod{\mathfrak{p}}.$$

Beweis. Die Notwendigkeit der Bedingung ist nach (4.4) klar. Um die Umkehrung zu zeigen, nehmen wir an, daß χ_α zum Block B_σ und χ_β zum Block $B_\tau \neq B_\sigma$ gehört. Wir verwenden jetzt eine Methode, die ganz analog der oben beim Beweis von (6C) verwendeten ist. Es sei $B_0 = B_\sigma \cup B_\tau$ die durch Zusammenwerfen der Blöcke B_σ und B_τ entstehende Menge von gewöhnlichen und modularen Charaktere. Geht man wieder von (5.5) aus und wendet man die Laplacesche Entwicklung an, wobei jetzt B_σ und B_τ durch die eine Menge B_0 ersetzt sind, so erhält man eine zu B_0 gehörige Unterdeterminante $\Delta_0 \not\equiv 0 \pmod{\mathfrak{p}}$. Durch weitere Entwicklung von Δ_0 sieht man, daß man annehmen darf, daß Δ_σ und Δ_σ in (5.7) komplementäre Unterdeterminanten von Δ_0 sind. Man darf also annehmen, daß die Indizes in Δ_0 gerade die Indizes in der Vereinigungsmenge $S_\sigma \cup S_\tau$ sind. Nach (6C) gibt es dann ein $r \in S_\sigma$ und ein $s \in S_\tau$ derart daß $\nu(G_r) = \nu(G_s) = d$ ist. Nach Voraussetzung ist nun $\omega_\alpha(G_r) \equiv \omega_\beta(G_r)$, $\omega_\alpha(G_s) \equiv \omega_\beta(G_s) \pmod{\mathfrak{p}}$. Daraus folgt nach (4.4), daß $\omega_i(G_r) \equiv \omega_j(G_r)$, $\omega_i(G_s) \equiv \omega_j(G_s) \pmod{\mathfrak{p}}$ für alle $\chi_i, \chi_j \in B_0 = B_\sigma \cup B_\tau$ ist. Drückt man jetzt in Δ_0 die Koeffizienten $\chi_i(G_r)$ und $\chi_i(G_s)$ in der r-ten und s-ten Spalte durch die entsprechenden ω_i aus, so erhält man genau wie oben die Ungleichung (6.4). Wegen $\nu(G_r) = \nu(G_s) = d$ ist dies ein Widerspruch, und damit ist (6F) bewiesen.

Schließlich zeigen wir noch:

(6G) *Ist eine Anzahl h von verschiedenen Charakteren χ_i vom Defekt 0 gegeben, so kann man h p-reguläre Elemente G_j vom Defekt 0 finden, derart daß*

$$\text{Det}\left(\chi_i(G_j)\right) \not\equiv 0 \pmod{\mathfrak{p}} \tag{6.5}$$

[11]) Vgl. § 13 der zweiten in [3]. genannten Arbeit.

ist, wo in der Determinante die Indizes i und j die h Werte durchlaufen, die zu den gegebenen Charakteren χ_i, bzw. zu den h gewählten Elementen G_j gehören.

Beweis. Nach (6E) ist jeder der m Charaktere χ_i der einzige gewöhnliche Charakter in seinem Block. Weiter folgt dann aus (5A), daß die h Charaktere χ_i in der Determinante (5.5) erscheinen. Durch Entwicklung nach den betreffenden Zeilen erhält man eine Unterdeterminante h-ten Grades, die die Bedingung (6.5) erfüllt. Wir haben zu zeigen, daß die hier auftretenden G sämtlich den Defekt 0 haben. Wäre eine der Zahlen $\nu(G_j)$ positiv, so folgte aus (2.3), daß $\chi_i(G_j) \equiv 0 \pmod{\mathfrak{p}}$ für alle unsere χ_i gilt, da $\omega_i(G_j)$ ganz algebraisch und g/x_i zu p teilerfremd ist. Dann wäre aber (6.5) unmöglich. Also haben alle G_j in (6.5) den Defekt 0.

§ 7. Zuordnung der Blöcke von $\mathfrak{G}$ zu Blöcken von Untergruppen

Es sei wie immer $\mathfrak{G}$ eine Gruppe endlicher Ordnung g, und p sei eine fest gewählte Primzahl. Wir betrachten Untergruppen $\mathfrak{T}$ und $\mathfrak{H}$ von $\mathfrak{G}$, die den folgenden Bedingungen genügen:

(*) Für jedes $T \in \mathfrak{T}$ gibt es eine p-Untergruppe $\mathfrak{P}_T$ von $\mathfrak{N}(T) \cap \mathfrak{N}(\mathfrak{T})$, derart daß $\mathfrak{C}(\mathfrak{P}_T) \subsetneqq \mathfrak{T}$ gilt.

(**) $$\mathfrak{T} \subsetneqq \mathfrak{H} \subsetneqq \mathfrak{N}(\mathfrak{T}).$$

Die Existenz derartiger Untergruppen $\mathfrak{T}$ ergibt sich aus der Bemerkung

(7A) *Ist $\mathfrak{Q}$ eine p-Untergruppe von $\mathfrak{G}$, so genügt $\mathfrak{T} = \mathfrak{C}(\mathfrak{Q})$ der Bedingung (*). Ist $\mathfrak{H}$ eine Untergruppe von $\mathfrak{N}(\mathfrak{Q})$, die $\mathfrak{T}$ enthält, so ist (**) erfüllt.*

In der Tat, ist T ein Element von $\mathfrak{T} = \mathfrak{C}(\mathfrak{Q})$, so kann man $\mathfrak{P}_T = \mathfrak{Q}$ wählen. Da $\mathfrak{T}$ invariant in $\mathfrak{N}(\mathfrak{Q})$ ist, ist $\mathfrak{N}(\mathfrak{Q}) \subsetneqq \mathfrak{N}(\mathfrak{T})$, also $\mathfrak{H} \subsetneqq \mathfrak{N}(\mathfrak{T})$.

Die Gruppen $\mathfrak{T}$ und $\mathfrak{H}$ in (7A) sind die einzigen Gruppen, für die die folgenden Betrachtungen angewendet werden.

(7B) *Es sei Ξ ein Körper der Charakteristik p, und es seien $\mathfrak{T}$ und $\mathfrak{H}$ Untergruppen von $\mathfrak{G}$, die den Bedingungen (*) und (**) genügen. Für jede Klasse $\mathfrak{K}_\alpha$ konjugierter Elemente von $\mathfrak{G}$ sei wie früher K_α das Element der Gruppenalgebra $\Gamma(\mathfrak{G}, \Xi)$, das die Summe der Elemente in $\mathfrak{K}_\alpha$ ist; es sei $K_\alpha^{(1)}$ die Summe der Elemente in $\mathfrak{K}_\alpha \cap \mathfrak{T}$[12]. Dann gibt es einen eindeutig bestimmten Homomorphismus H der Klassenalgebra $Z(\mathfrak{G}, \Xi)$ von $\mathfrak{G}$ in die Klassenalgebra $Z(\mathfrak{H}, \Xi)$ von $\mathfrak{H}$, der K_α auf $K_\alpha^{(1)}$ abbildet.*

Beweis. Offenbar besteht $\mathfrak{K}_\alpha \cap \mathfrak{T}$ aus vollen Klassen von $\mathfrak{H}$, und daher kann $K_\alpha^{(1)}$ als Element von $Z(\mathfrak{H}, \Xi)$ betrachtet werden. Man setze $K_\alpha = K_\alpha^{(1)} + K_\alpha^{(2)}$, so daß also $K_\alpha^{(2)}$ die Summe der nicht in $\mathfrak{T}$ liegenden Elemente von $\mathfrak{K}_\alpha$ ist. Ist die Multiplikation der K_α durch die Formel (2.2) gegeben, so haben wir

$$(K_\alpha^{(1)} + K_\alpha^{(2)})(K_\beta^{(1)} + K_\beta^{(2)}) = \sum_\gamma a_{\alpha\beta\gamma}(K_\gamma^{(1)} + K_\gamma^{(2)}).$$

[12]) Gibt es kein derartiges Element, so ist $K_\alpha^{(1)} = 0$ zu setzen.

Man betrachte hier beide Seiten als Summen von Elementen von $\mathfrak{G}$ und vergleiche die Koeffizienten der in $\mathfrak{T}$ liegenden Elemente. Offenbar besteht $K_\alpha^{(1)} K_\beta^{(1)}$ ausschließlich aus solchen Elementen, während in $K_\alpha^{(1)} K_{\beta_1}^{(2)}$ und in $K_\alpha^{(2)} K_\beta^{(1)}$ nur außerhalb $\mathfrak{T}$ liegende Elemente auftreten. Können wir also zeigen, daß jedes in $K_\alpha^{(2)} K_\beta^{(2)}$ erscheinende Element von $\mathfrak{T}$ einen durch p teilbaren Koeffizienten hat, so ergibt der Koeffizientenvergleich, daß in $Z(\mathfrak{H}, \Xi)$ die Gleichung gilt

$$K_\alpha^{(1)} K_\beta^{(1)} = \sum_\gamma a_{\alpha\beta\gamma} K_\gamma^{(1)}. \tag{7.1}$$

Da die K_α eine Basis von $Z(\mathfrak{G}, \Xi)$ bilden, zeigt Vergleich mit (2.2), daß die durch $K_\alpha \to K_\alpha^{(1)}$ definierte lineare Abbildung H von $Z(\mathfrak{G}, \Xi)$ in $Z(\mathfrak{H}, \Xi)$ ein Homomorphismus ist.

Die Vielfachheit N, mit der ein Element T von $\mathfrak{T}$ in $K_\alpha^{(2)} K_\beta^{(2)}$ auftritt, ist offenbar die Anzahl der geordneten Paare (X, Y) von Elementen von $\mathfrak{G}$, die den Bedingungen genügen

$$XY = T; \quad X \in \mathfrak{K}_\alpha; \quad X \notin \mathfrak{T}; \quad Y \in \mathfrak{K}_\beta; \quad Y \notin \mathfrak{T}.$$

Ist nun P ein Element der Gruppe $\mathfrak{P}_T$ in (*), so erfüllt mit (X, Y) auch das Paar $(P^{-1}XP, P^{-1}YP)$ die gleichen Bedingungen. Nennen wir also zwei Paare (X, Y) und (X', Y') äquivalent, wenn es ein P in $\mathfrak{P}_T$ mit $X' = P^{-1}XP$, $Y' = P^{-1}YP$ gibt, so sind unsere N Paare auf Äquivalenzklassen verteilt. Die (X, Y) enthaltende Äquivalenzklasse besteht offenbar aus $(\mathfrak{P}_T : \mathfrak{P}_T \cap \mathfrak{N}(X) \cap \mathfrak{N}(Y))$ verschiedenen Paaren. Dieser Gruppenindex ist eine Potenz von p. Wäre er 1, so wäre $\mathfrak{P}_T \subseteq \mathfrak{N}(X)$, also $X \in \mathfrak{C}(\mathfrak{P}_T)$, daher wegen (*) $X \in \mathfrak{T}$. Da dies einen Widerspruch darstellt, ist der Gruppenindex und dann auch N durch p teilbar. Das ist aber, was wir zu zeigen hatten, und damit ist (7B) bewiesen.

Als Zusatz haben wir

(7C) *Ist* $\mathfrak{U}$ *der Kern des Homomorphismus* H *in* (7B), *so besteht eine Basis von* $\mathfrak{U}$ *aus denjenigen* K_α, *für die* $\mathfrak{K}_\alpha$ *keine Elemente von* $\mathfrak{T}$ *enthält. Eine Basis des Bildes* $\mathfrak{V}$ *von* $Z(\mathfrak{G}, \Xi)$ *in* $Z(\mathfrak{H}, \Xi)$ *besteht aus den von* 0 *verschiedenen* $K_\alpha^{(1)}$.

Dabei hat man nur zu beachten, daß die verschiedenen $K_\alpha^{(1)} \neq 0$ sicher linear unabhängig sind, da sich zwei verschiedene von ihnen aus gänzlich verschiedenen Elementen zusammensetzen.

In der eben eingeführten Bezeichnung haben wir

$$Z(\mathfrak{G}, \Xi)/\mathfrak{U} \cong \mathfrak{V} \subseteq Z(\mathfrak{H}, \Xi). \tag{7.2}$$

Wir wählen jetzt für Ξ den Körper Ω^* aus § 2. Ist ψ ein linearer Charakter von $Z(\mathfrak{G}, \Omega^*)$, dessen Kern das Ideal $\mathfrak{U}$ enthält, so kann man ψ als linearen Charakter von $Z(\mathfrak{G}, \Omega^*)/\mathfrak{U}$ betrachten. Durch Anwendung des Isomorphismus in (7.2) erhält man daraus einen linearen Charakter der Teilalgebra $\mathfrak{V}$ von $Z(\mathfrak{H}, \Omega^*)$, und dieser kann dann zu einem linearen Charakter $\tilde{\psi}$ von $Z(\mathfrak{H}, \Omega^*)$

erweitert werden[13]). Geht man umgekehrt von einem linearen Charakter $\tilde{\psi}$ von $Z(\mathfrak{H}, \Omega^*)$ aus, so erhält man durch Beschränkung einen linearen Charakter von $\mathfrak{V}$. Mit Hilfe von (7.2) ergibt sich daraus ein linearer Charakter von $Z(\mathfrak{G}, \Omega^*)/\mathfrak{U}$, und dieser kann als linearer Charakter von $Z(\mathfrak{G}, \Omega^*)$ gedeutet werden, dessen Kern $\mathfrak{U}$ enthält. In dieser Weise erkennen wir:

(7D) *Die linearen Charaktere $\tilde{\psi}$ von $Z(\mathfrak{H}, \Omega^*)$ entsprechen eineindeutig den linearen Charakteren ψ von $Z(\mathfrak{G}, \Omega^*)$, deren Kern das Ideal $\mathfrak{U}$ in* (7C) *(für $\Xi = \Omega^*$) enthält. Für zugeordnete Charaktere $\tilde{\psi}$ und ψ gilt*

$$\psi(K_\alpha) = \tilde{\psi}(K_\alpha^{(1)}) = \sum_\varrho \tilde{\psi}(L_\varrho), \tag{7.3}$$

wo auf der rechten Seite $\mathfrak{L}_\varrho$ über diejenigen Klassen von $\mathfrak{H}$ läuft, die in $\mathfrak{K}_\alpha \cap \mathfrak{T}$ liegen, und wo L_ϱ die Summe der Elemente in $\mathfrak{L}_\varrho$ ist.

Wie wir in § 4 gesehen haben, entsprechen die linearen Charaktere der Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$ den Blöcken von $\mathfrak{G}$. Ebenso entsprechen die linearen Charaktere der Klassenalgebra $Z(\mathfrak{H}, \Omega^*)$ den Blöcken von $\mathfrak{H}$. Daher folgt:

(7E) *Auf Grund der Zuordnung $\tilde{\psi} \to \psi$ in* (7D) *entspricht jedem Block $\tilde{B}$ von $\mathfrak{H}$ ein Block B von $\mathfrak{G}$. Dabei ist $\tilde{\psi}$ der dem Block $\tilde{B}$ zugeordnete lineare Charakter der Klassenalgebra $Z(\mathfrak{H}, \Omega^*)$ und ψ der dem Block B zugeordnete lineare Charakter von $Z(\mathfrak{G}, \Omega^*)$.*

Ist χ_i ein zum Block B gehöriger Charakter von $\mathfrak{G}$, und ist ω_i wie immer durch (2.3) definiert, so ist $\psi = \omega_i^*$. Ähnlich kann man $\tilde{\psi}$ durch einen zu $\tilde{B}$ gehörigen Charakter $\tilde{\chi}_i$ von $\mathfrak{H}$ ausdrücken. Setzt man dies in (7.3) ein, so erhält man nach einfacher Umformung das Ergebnis:

(7F) *Ist χ_i vom Grad x_i ein irreduzibler Charakter des Blocks B von $\mathfrak{G}$ und $\tilde{\chi}_j$ vom Grad $\tilde{x}_j$ ein irreduzibler Charakter des Blocks $\tilde{B}$ von $\mathfrak{H}$ in* (7E), *so gilt für jedes Element G von $\mathfrak{G}$ die Kongruenz*[14])

$$\chi_i(G) \equiv \frac{n(G)}{g} \frac{x_i}{\tilde{x}_j} \sum \tilde{\chi}_j(H) \pmod{\mathfrak{p}_\nu \mathfrak{p}^{\nu(G) - \nu(g/x_i)}}, \tag{7.4}$$

wo H über alle Elemente von $\mathfrak{T}$ läuft, die in $\mathfrak{G}$ zu G konjugiert sind. Gibt es keine derartigen Elemente, so ist die Summe gleich 0 zu setzen.

§ 8. Die Defektgruppe eines Blockes B

Es sei B ein Block von $\mathfrak{G}$ vom Defekt d. Es sei ψ der dem Block B zugeordnete lineare Charakter der Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$. Aus (6B) folgt, daß es Klassen $\mathfrak{K}_\alpha$ vom Defekt d gibt, für die $\psi(K_\alpha) \neq 0$ ist. Ist G ein Element einer derartigen Klasse $\mathfrak{K}_\alpha$ und ist $\mathfrak{D}$ eine p-Sylowgruppe des Normalisators $\mathfrak{N}(G)$, so nennen wir $\mathfrak{D}$ eine *Defektgruppe* des Blocks B. Es ist dann klar,

[13]) Vgl. § 14, (14A).

[14]) Die Bedeutung dieser Kongruenz ist natürlich folgendermaßen gemeint: Ist δ die Differenz der linken und rechten Seite von (7.4), so ist

$$\nu(\delta) > \nu(G) - \nu(g/x_i).$$

daß jede zu $\mathfrak{D}$ konjugierte Gruppe ebenfalls eine Defektgruppe von B ist. Da $\nu(G)=d$ ist, hat jede Defektgruppe die Ordnung p^d.

(8A) *Ist ψ der dem Block B zugeordnete lineare Charakter von $Z(\mathfrak{G},\Omega^*)$, und ist $\psi(K_\beta)\neq 0$ für eine Klasse $\mathfrak{K}_\beta$, so enthält der Normalisator $\mathfrak{N}(G_\beta)$ eines jeden Elementes G_β von $\mathfrak{K}_\beta$ eine Defektgruppe von B.*

Beweis. Es sei wie oben die Klasse $\mathfrak{K}_\alpha$ derart gewählt, daß $\psi(K_\alpha)\neq 0$ ist, und daß $\mathfrak{K}_\alpha$ den Defekt d hat. Ist die Multiplikation der K_i wie immer durch (2.2) definiert, so hat man

$$\psi(K_\alpha)\psi(K_\beta)=\sum_\gamma a_{\alpha\beta\gamma}\psi(K_\gamma).$$

Ist also $\psi(K_\beta)\neq 0$, so muß es einen Index γ geben, für den

$$a_{\alpha\beta\gamma}\not\equiv 0 \pmod{p}; \qquad \psi(K_\gamma)\neq 0 \tag{8.1}$$

ist. Es sei G_γ ein Element von $\mathfrak{K}_\gamma$ und Ω von der Ordnung p^r eine p-Sylowgruppe von $\mathfrak{N}(G_\gamma)$. Dann muß also $\nu(G_\gamma)=r$ sein. Aus der zweiten Bedingung (8.1) folgt nach (6B), daß $d\leq r$ ist.

Jetzt nehme man $\mathfrak{T}=\mathfrak{C}(\Omega)$, $\mathfrak{H}=\mathfrak{N}(\Omega)$. Wie wir in (7A) gesehen haben, sind dann die Bedingungen (*) und (**) erfüllt, und daher gelten die Resultate von § 7. Nach (7.1) ist also

$$K_\alpha^{(1)}K_\beta^{(1)}=\sum_\gamma a_{\alpha\beta\gamma}K_\gamma^{(1)}.$$

Nun ist $G_\gamma\in\mathfrak{T}\cap\mathfrak{K}_\gamma$ also $K_\gamma^{(1)}\neq 0$. Wegen der linearen Unabhängigkeit der von 0 verschiedenen $K_i^{(1)}$ müssen $K_\alpha^{(1)}$ und $K_\beta^{(1)}$ beide von 0 verschieden sein. Daher gibt es ein Element G_α in $\mathfrak{K}_\alpha$ und ein Element G_β in $\mathfrak{K}_\beta$, die zu $\mathfrak{T}$ gehören. Dann ist $\Omega\leq\mathfrak{N}(G_\alpha)$ und $\Omega\leq\mathfrak{N}(G_\beta)$. Da $\nu(G_\alpha)=d$ ist und Ω die Ordnung p^r hat, muß $d\geq r$ und also $d=r$ sein. Folglich ist Ω eine Sylowgruppe von $\mathfrak{N}(G_\alpha)$ und, da $\psi(K_\alpha)\neq 0$ war, ist Ω eine Defektgruppe von B. Damit haben wir gezeigt, daß $\mathfrak{N}(G_\beta)$ eine Defektgruppe von B enthält. Das gleiche gilt dann für den Normalisator jedes zu G_β konjugierten Elements, wie zu beweisen war.

Ist $\mathfrak{D}$ eine beliebige Defektgruppe von B, so gibt es nach Definition eine Klasse $\mathfrak{K}_\beta$ mit $\psi(K_\beta)\neq 0$, derart daß $\mathfrak{D}$ Sylowgruppe des Normalisators $\mathfrak{N}(G)$ eines Elements G von $\mathfrak{K}_\beta$ ist. Dann muß also $\nu(G)=d$ sein. Wählt man eine feste Klasse $\mathfrak{K}_\alpha$ vom Defekt d mit $\psi(K_\alpha)\neq 0$ und wendet die Methode im Beweis von (8A) an, so sieht man, daß es eine Sylowgruppe Ω eines Normalisators $\mathfrak{N}(G_\alpha)$ mit $G_\alpha\in\mathfrak{K}_\alpha$ gibt, derart daß Ω zum Normalisator von Elementen G_β von $\mathfrak{K}_\beta$ gehört. Dann muß Ω die Ordnung p^d haben und also Sylowgruppe von $\mathfrak{N}(G_\beta)$ sein. Da G_β und G als Elemente der Klasse $\mathfrak{K}_\beta$ konjugiert sind, sind die Sylowgruppen Ω und $\mathfrak{D}$ ihrer Normalisatoren konjugiert. Damit haben wir gezeigt, daß eine beliebige Defektgruppe $\mathfrak{D}$ zu den Sylowgruppen der Normalisatoren der Elemente einer festen Klasse $\mathfrak{K}_\alpha$ konjugiert sind. Wir haben also bewiesen:

(8B) *Je zwei Defektgruppen eines Blocks B sind konjugiert.*

Wir zeigen weiter:

(8C) *Ist B ein Block von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$, und setzt man $\mathfrak{T}=\mathfrak{C}(\mathfrak{D})$, $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$, so gibt es einen Block $\tilde{B}$ von $\mathfrak{H}$, dem B im Sinn von (7E) zugeordnet ist.*

Beweis. Es sei ψ der zu B gehörige lineare Charakter der Klassenalgebra $Z(\mathfrak{G},\Omega^*)$. Ist $\psi(K_\beta)\neq 0$ für eine Klasse $\mathfrak{K}_\beta$, so folgt aus (8A), daß $\mathfrak{K}_\beta$ Elemente G enthält, für die $\mathfrak{D}\leqq\mathfrak{N}(G)$ ist. Dann ist $G\in\mathfrak{C}(\mathfrak{D})=\mathfrak{T}$, also $\mathfrak{K}_\beta\cap\mathfrak{T}$ nicht leer und $K_\beta^{(1)}\neq 0$. Ist daher $K_\beta^{(1)}=0$, so muß $\psi(K_\beta)$ verschwinden. In der Bezeichnung von (7C) enthält also der Kern von ψ das Ideal $\mathfrak{U}$. Jetzt zeigt (7D), daß es einen linearen Charakter $\tilde{\psi}$ von $Z(\mathfrak{H},\Omega^*)$ gibt, dem ψ zugeordnet ist, und daraus ergibt sich die Behauptung.

(8D) *Sind $\mathfrak{T}$ und $\mathfrak{H}$ zwei beliebige Untergruppen von $\mathfrak{G}$, die die Bedingungen (*) und (**) von § 7 erfüllen, und ist $\tilde{B}$ ein Block von $\mathfrak{H}$, dem der Block B von $\mathfrak{G}$ im Sinn von (7E) entspricht, so ist die Defektgruppe $\tilde{\mathfrak{D}}$ von $\tilde{B}$ in einer Defektgruppe $\mathfrak{D}$ von B enthalten. Insbesondere ist der Defekt $\tilde{d}$ von $\tilde{B}$ höchstens gleich dem Defekt d von B.*

Beweis. Nach Definition der Defektgruppe gibt es eine Klasse $\mathfrak{K}_\alpha$ vom Defekt d, für die $\psi(K_\alpha)\neq 0$ für den zu B gehörigen linearen Charakter ψ von $Z(\mathfrak{G},\Omega^*)$ ist. Der zu $\tilde{B}$ gehörige lineare Charakter $\tilde{\psi}$ von $Z(\mathfrak{H},\Omega^*)$ ist mit ψ durch die Formel (7.3) verbunden. Daraus folgt, daß es eine in $\mathfrak{K}_\alpha\cap\mathfrak{T}$ enthaltene Klasse $\mathfrak{L}_\varrho$ von $\mathfrak{H}$ gibt, für die $\tilde{\psi}(L_\varrho)\neq 0$ ist. Ist G ein Element von $\mathfrak{L}_\varrho$, so zeigt Anwendung von (8A) auf die Gruppe $\mathfrak{H}$, daß der Normalisator $\mathfrak{N}(G)\cap\mathfrak{H}$ von G in $\mathfrak{H}$ eine Defektgruppe $\tilde{\mathfrak{D}}$ von $\tilde{B}$ enthält. Dann gehört $\tilde{\mathfrak{D}}$ zu einer Sylowgruppe $\mathfrak{D}$ von $\mathfrak{N}(G)$, und da $G\in\mathfrak{K}_\alpha$ ist, ist $\mathfrak{D}$ eine Defektgruppe von B. Der letzte Teil der Behauptung folgt aus $(\mathfrak{D}:1)=p^d$, $(\tilde{\mathfrak{D}}:1)=p^{\tilde{d}}$.

§ 9. Gruppen mit invarianten p-Untergruppen

Wir unterbrechen jetzt unsere allgemeinen Betrachtungen und behandeln den speziellen Fall einer Gruppe $\mathfrak{G}$, die eine invariante Untergruppe $\mathfrak{M}$ enthält. Speziell sind wir an dem Fall interessiert, daß $\mathfrak{M}$ eine p-Gruppe ist, da die Ergebnisse später auf die Gruppe $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$ in (8C) an Stelle von $\mathfrak{G}$ angewendet werden sollen. Die Untersuchung ist dabei etwas weitergeführt, als es für unsere Zwecke unbedingt nötig ist.

Wir bezeichnen die Faktorgruppe von $\mathfrak{G}$ nach der invarianten Untergruppe $\mathfrak{M}$ mit $\overline{\mathfrak{G}}$;

$$\overline{\mathfrak{G}}=\mathfrak{G}/\mathfrak{M}. \tag{9.1}$$

Ist G ein Element von $\mathfrak{G}$, so sei stets $\overline{G}=G\mathfrak{M}$ das entsprechende Element von $\overline{\mathfrak{G}}$.

(9A) *Für $G\in\mathfrak{G}$ ist $\mathfrak{N}(G)\mathfrak{M}/\mathfrak{M}$ eine Untergruppe des Normalisators $\mathfrak{N}(\overline{G})$ von $\overline{G}$ in $\overline{\mathfrak{G}}$. Die Ordnung $n(\overline{G})$ von $\mathfrak{N}(\overline{G})$ ist teilbar durch $n(G)/z$, wo z die Ordnung von $\mathfrak{N}(G)\cap\mathfrak{M}=\mathfrak{Z}$ ist.*

Der erste Teil der Behauptung ist klar. Der zweite Teil folgt daraus, daß $\mathfrak{N}(G)\mathfrak{M}/\mathfrak{M} \simeq \mathfrak{N}(G)/\mathfrak{Z}$ die Ordnung $n(G)/z$ hat.

Jeder Charakter von $\overline{\mathfrak{G}}$ kann als Charakter von $\mathfrak{G}$ aufgefaßt werden, dessen Kern die Gruppe $\mathfrak{M}$ enthält. In diesem Sinn behaupten wir:

(9B) *Jeder Block $\overline{B}$ von $\overline{\mathfrak{G}}$ ist in einem Block B von $\mathfrak{G}$ enthalten.*

Beweis. Wir setzen $m = (\mathfrak{M}:1)$. Ist χ_i ein gewöhnlicher irreduzibler Charakter von $\overline{\mathfrak{G}}$, so definieren wir $\overline{\omega}_i$ in Analogie mit (2.3) durch

$$\overline{\omega}_i(\overline{G}) = \frac{g}{m\,\bar{n}(\overline{G})}\,\frac{\chi_i(\overline{G})}{x_i}, \qquad \left(x_i = \chi_i(1)\right).$$

Betrachten wir χ_i jetzt als irreduziblen Charakter von $\mathfrak{G}$ und hat ω_i die alte Bedeutung (2.3), so ist also

$$\omega_i(G) = \frac{m\,\bar{n}(\overline{G})}{n(G)}\,\overline{\omega}_i(\overline{G}), \tag{9.2}$$

da wir $\chi_i(G) = \chi_i(\overline{G})$ zu setzen haben. Aus (9A) folgt, daß $\bar{n}(\overline{G})\,m/n(G)$ ganz ist, da ja z ein Teiler von m ist.

Gehören nun zwei irreduzible Charaktere χ_i und χ_j von $\overline{\mathfrak{G}}$ zum selben Block $\overline{B}$, so ist nach (4.4) $\overline{\omega}_i(\overline{G}) \equiv \overline{\omega}_j(\overline{G}) \pmod{\mathfrak{p}}$ für jedes G in $\mathfrak{G}$. Dann zeigt (9.2), daß $\omega_i(G) \equiv \omega_j(G) \pmod{\mathfrak{p}}$ für jedes G in $\mathfrak{G}$ ist. Daraus folgt, daß χ_i und χ_j als Charaktere von $\mathfrak{G}$ zum selben Block B gehören. Ist $\mathfrak{X}_i$ die Darstellung von $\overline{\mathfrak{G}}$ mit dem Charakter χ_i, so kann man die irreduziblen modularen Darstellungen von $\overline{B}$ als die in den Darstellungen $\mathfrak{X}_i^*$ für $\mathfrak{X}_i \in \overline{B}$ wirklich auftretenden irreduziblen Bestandteile $\mathfrak{F}_j$ definieren; vgl. (4D). Faßt man jetzt alle diese Darstellungen als Darstellungen von $\mathfrak{G}$ auf, so sieht man, daß alle $\mathfrak{F}_j \in \overline{B}$ zu B gehören, und damit ist (9B) bewiesen.

(9C) *Hat B in* (9B) *die Defektgruppe $\mathfrak{D}$, so enthält $\mathfrak{D}$ eine p-Sylowgruppe von $\mathfrak{M}$. Die Untergruppe $\mathfrak{M}\mathfrak{D}/\mathfrak{M}$ von $\overline{\mathfrak{G}}$ enthält eine Defektgruppe $\overline{\mathfrak{D}}$ von $\overline{B}$, und daher ist der Defekt $\bar{d}$ von $\overline{B}$ höchstens gleich $d - \nu(m)$.*

Beweis. Ist χ_i ein zu $\overline{B}$ und also zu B gehöriger Charakter, und ist ψ der B entsprechende lineare Charakter der Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$, so ist $\omega_i^* = \psi$, § 4. Nach Definition der Defektgruppe $\mathfrak{D}$ gibt es Elemente $G \in \mathfrak{G}$ vom Defekt d, derart daß $\psi(K_\alpha) \not\equiv 0$ ist, wobei $\mathfrak{K}_\alpha$ die G enthaltende Klasse ist. Dann ist $\omega_i(G) \not\equiv 0 \pmod{\mathfrak{p}}$. Aus (9.2) folgt dann

$$\nu(m) + \nu\left(\bar{n}(\overline{G})\right) = \nu\left(n(G)\right) = d; \quad \overline{\omega}_i(\overline{G}) \not\equiv 0 \pmod{\mathfrak{p}}. \tag{9.3}$$

Da nach (9A) $\nu\left(\bar{n}(\overline{G})\right) \geq \nu\left(n(G)\right) - \nu(z) \geq \nu\left(n(G)\right) - \nu(m)$ ist, muß $\nu(z) = \nu(m)$ sein. Jede Sylowgruppe $\mathfrak{P}_0$ von $\mathfrak{Z}$ ist daher auch Sylowgruppe von $\mathfrak{M}$. Wegen $\mathfrak{Z} \subseteq \mathfrak{N}(G)$ kann man annehmen, daß $\mathfrak{P}_0$ in einer Sylowgruppe von $\mathfrak{N}(G)$, d.h. in einer Defektgruppe $\mathfrak{D}$ von B enthalten ist. Offenbar enthält dann auch jede zu $\mathfrak{D}$ konjugierte Gruppe eine Sylowgruppe von $\mathfrak{M}$. Aus dem zweiten Teil von (9.3) ergibt sich ferner nach (8A), daß jede Sylowgruppe von $\mathfrak{N}(\overline{G})$ eine Defektgruppe $\overline{\mathfrak{D}}$ von $\overline{B}$ enthält. Nach dem ersten Teil von (9.3) hat eine solche Sylowgruppe die Ordnung $p^{d-\nu(m)}$. Offenbar hat die Untergruppe

$\mathfrak{MD}/\mathfrak{M}$ von $\mathfrak{N}(\overline{G})$ diese Ordnung, ist also Sylowgruppe, und daher kann man $\mathfrak{D} \subseteq \mathfrak{MD}/\mathfrak{M}$ annehmen, wie behauptet war.

Von jetzt an setzen wir voraus, daß die invariante Untergruppe von $\mathfrak{G}$ eine p-Gruppe ist.

(9D) *Enthält $\mathfrak{G}$ eine invariante Untergruppe $\mathfrak{M}$ der Ordnung p^μ, so ist $\mathfrak{M}$ in dem Kern jeder irreduziblen modularen Darstellung $\mathfrak{F}$ von $\mathfrak{G}$ enthalten.*

Beweis. Beschränkt man in der Darstellung $\mathfrak{F}$ das Argument auf $\mathfrak{M}$, so erhält man eine Darstellung von $\mathfrak{M}$, die nach einem Satz von A. H. Clifford[15]) vollständig reduzibel ist. Nun ist die einzige irreduzible modulare Darstellung einer p-Gruppe die Einsdarstellung [z. B. nach (3B)]. Daher ist $\mathfrak{F}(M)$ die Identität für jedes $M \in \mathfrak{M}$, und daraus ergibt sich die Behauptung.

(9E) *Enthält die Gruppe $\mathfrak{G}$ eine invariante Untergruppe $\mathfrak{M}$ der Ordnung p^μ, so enthält jeder Block B von $\mathfrak{G}$ vom Defekt d mindestens einen Block $\overline{B}$ von $\overline{\mathfrak{G}}$ vom Defekt $\bar{d} = d - \mu$. Alle in B enthaltenen Blöcke von $\overline{\mathfrak{G}}$ haben höchstens den Defekt $d - \mu$.*

Beweis. Nach (6A) kann man in B eine irreduzible modulare Darstellung $\mathfrak{F}$ finden, deren Grad f die Primzahl p genau mit dem Exponenten $\nu(g) - d$ enthält. Nach (9D) kann man $\mathfrak{F}$ als Darstellung von $\overline{\mathfrak{G}} = \mathfrak{G}/\mathfrak{M}$ betrachten. Gehört dann $\mathfrak{F}$ zu dem Block $\overline{B}$ vom Defekt $\bar{d}$, so muß f durch $p^{\nu(g)-\mu-\bar{d}}$ teilbar sein, da die Ordnung von $\overline{\mathfrak{G}}$ die Primzahl p mit dem Exponenten $\nu(g) - \mu$ enthält. Dies liefert $d \leq \bar{d} + \mu$.

Andererseits ist nach (9C) der Defekt eines beliebigen in B enthaltenen Blocks von $\overline{\mathfrak{G}}$ höchstens gleich $d - \mu$. Dies liefert die letzte Behauptung in (9E). Für $\bar{d}$ ergibt sich jetzt, daß $\bar{d} = d - \mu$ ist, und damit ist auch der erste Teil von (9E) bewiesen.

Wir heben noch den folgenden speziellen Fall unserer Ergebnisse hervor:

(9F) *Enthält $\mathfrak{G}$ eine invariante Untergruppe $\mathfrak{M}$ der Ordnung p^μ, so ist $\mathfrak{M}$ in der Defektgruppe $\mathfrak{D}$ jedes Blocks B von $\mathfrak{G}$ enthalten, und daher ist der Defekt d von B mindestens gleich μ.*

Da nämlich der Block B von $\mathfrak{G}$ nach (9E) einen Block $\overline{\mathfrak{B}}$ von $\overline{\mathfrak{G}}$ enthält, kann man (9C) anwenden. Da $\mathfrak{M}$ seine eigene Sylowgruppe ist, ergibt sich $\mathfrak{M} \subseteq \mathfrak{D}$ und daraus $\mu \leq d$.

§ 10. Der Hauptsatz

Wir brauchen den folgenden Hilfssatz:

(10A) *Es sei $\mathfrak{D}$ eine p-Untergruppe von $\mathfrak{G}$, und es werde $\mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$ gesetzt. Gibt es in einer Klasse $\mathfrak{K}_i$ von $\mathfrak{G}$ Elemente G, für die $\mathfrak{D}$ Sylowgruppe des Normalisators $\mathfrak{N}(G)$ ist, so enthält $\mathfrak{K}_i$ genau eine aus Elementen von $\mathfrak{T}$ bestehende Klasse $\mathfrak{L}_j$ von $\mathfrak{H}$. Ist H ein Element von $\mathfrak{H}$ und ist $\mathfrak{D}$ Sylowgruppe des Normalisators $\mathfrak{N}(H) \cap \mathfrak{H}$ von H in $\mathfrak{H}$, so ist $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(H)$, und daher kann die erste Behauptung auf die Klasse $\mathfrak{K}_i$ von $\mathfrak{G}$ angewendet werden, in der H liegt.*

[15]) Clifford, A. H.: Ann. of Math. **38**, 533—550 (1937). Vgl. auch H. Weyl, The classical groups. Princeton 1939.

Beweis. Ist $\mathfrak{D} \subsetneq \mathfrak{N}(G)$, so ist $G \in \mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, und daher enthält die Klasse $\mathfrak{K}_i$ von G Elemente von $\mathfrak{T}$. Ist G' ein zweites Element von $\mathfrak{K}_i \cap \mathfrak{T}$, so kann man $G' = X^{-1}GX$ mit $X \in \mathfrak{G}$ setzen. Wegen $G' \in \mathfrak{T}$, ist $\mathfrak{D} \subseteq \mathfrak{N}(G')$, und da $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(G)$ ist, so ist $\mathfrak{D}$ ebenfalls Sylowgruppe von $\mathfrak{N}(G')$. Da $X^{-1}\mathfrak{D}X$ zu $\mathfrak{N}(X^{-1}GX) = \mathfrak{N}(G')$ gehört, ist $X^{-1}\mathfrak{D}X$ auch Sylowgruppe von $\mathfrak{N}(G')$ und daher zu $\mathfrak{D}$ in $\mathfrak{N}(G')$ konjugiert. Man kann also ein Element Y in $\mathfrak{N}(G')$ finden, so daß $X^{-1}\mathfrak{D}X = Y^{-1}\mathfrak{D}Y$ ist. Dann gehört XY^{-1} zu $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$, und da XY^{-1} das Element G in G' transformiert, so gehören G und G' zu derselben Klasse von $\mathfrak{H}$. Damit ist der erste Teil von (10A) bewiesen.

Es sei H ein Element von $\mathfrak{H}$, für das $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(H) \cap \mathfrak{H}$ ist, und es sei $\mathfrak{D}_1$ eine $\mathfrak{D}$ enthaltende Sylowgruppe von $\mathfrak{N}(H)$. Wäre $\mathfrak{D}_1 > \mathfrak{D}$, so gäbe es nach bekannten Sätzen über p-Gruppen Elemente M in $\mathfrak{D}_1$, die zu $\mathfrak{N}(\mathfrak{D}) = \mathfrak{H}$, aber nicht zu $\mathfrak{D}$ gehören. Dann wäre $\{M, \mathfrak{D}\}$ eine von $\mathfrak{D}$ verschiedene p-Gruppe, die in $\mathfrak{N}(H) \cap \mathfrak{H}$ enthalten ist. Dies ist ein Widerspruch, da dann $\mathfrak{D}$ nicht Sylowgruppe von $\mathfrak{N}(H) \cap \mathfrak{H}$ wäre. Also ist $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(H)$, wie behauptet war.

Jetzt sind wir in der Lage, den folgenden Satz zu beweisen, der das Hauptergebnis dieser Arbeit darstellt:

(10B) *Es sei $\mathfrak{D}$ eine Untergruppe von $\mathfrak{G}$, deren Ordnung eine Potenz p^d von p ist. Man setze $\mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$. Dann besteht eine eineindeutige Zuordnung zwischen den Blöcken B von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$ und den Blöcken $\tilde{B}$ von $\mathfrak{H}$ vom Defekt d. Dabei ist B dem Block $\tilde{B}$ im Sinn von § 7 zugeordnet, so daß die Formeln* (7.3) *und* (7.4) *gelten. Die Blöcke $\tilde{B}$ von $\mathfrak{H}$ vom Defekt d besitzen $\mathfrak{D}$ als Defektgruppe.*

Beweis. Ist B ein Block von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$, so gibt es nach (8C) einen Block $\tilde{B}$ von $\mathfrak{H}$, dem B im Sinn von § 7 zugeordnet ist. Aus (8D) ergibt sich, daß der Defekt $\tilde{d}$ von $\tilde{B}$ höchstens gleich dem Defekt von B ist. Da aber B eine Defektgruppe $\mathfrak{D}$ von der Ordnung p^d hat, hat B den Defekt d; es ist also $\tilde{d} \leq d$. Die Gruppe $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$ hat eine invariante Untergruppe $\mathfrak{D}$ der Ordnung p^d. Wendet man also (9F) auf die Gruppe $\mathfrak{H}$ an Stelle von $\mathfrak{G}$ an, so sieht man, daß der Defekt $\tilde{d}$ von $\tilde{B}$ mindestens gleich d ist. Damit ist gezeigt, daß $\tilde{d} = d$ ist. Der eben verwendete Hilfssatz (9F) zeigt auch noch, daß jeder Block von $\mathfrak{H}$ vom Defekt d die Defektgruppe $\mathfrak{D}$ besitzt.

Es sei $\tilde{B}$ ein beliebiger Block von $\mathfrak{H}$ mit der Defektgruppe $\mathfrak{D}$. Ist $\tilde{\psi}$ der zu $\tilde{B}$ gehörige lineare Charakter der Klassenalgebra $Z(\mathfrak{H}, \Omega^*)$, so kann man also eine Klasse $\mathfrak{L}_j$ von $\mathfrak{H}$ und ein Element H in $\mathfrak{L}_j$ so wählen, daß $\tilde{\psi}(L_j) \neq 0$ ist, und daß $\mathfrak{D}$ eine Sylowgruppe des Normalisators $\mathfrak{N}(H) \cap \mathfrak{H}$ von H in $\mathfrak{H}$ ist. Aus (10A) folgt dann, daß $\mathfrak{D}$ dann auch Sylowgruppe von $\mathfrak{N}(H)$ selbst ist, und weiter, daß wenn H in $\mathfrak{G}$ zur Klasse $\mathfrak{K}_i$ gehört, man $\mathfrak{L}_j = \mathfrak{K}_i \cap \mathfrak{T}$ hat.

Für gegebenes $\tilde{\psi}$ definieren die Formeln (7.3) einen linearen Charakter ψ der Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$. Dieser gehört zu dem Block B von $\mathfrak{G}$, der dem Block $\tilde{B}$ von $\mathfrak{H}$ nach § 7 zugeordnet ist. Wendet man die Formeln (7.3)

auf die Klasse $\mathfrak{K}_i$ an, so hat man nach dem eben Bewiesenen

$$\psi(K_i) = \tilde{\psi}(L_j). \tag{10.1}$$

Da $\tilde{\psi}(L_j) \neq 0$ war, ist $\psi(K_i) \neq 0$. Dann folgt aber aus (8A), daß $\mathfrak{N}(H)$ eine Defektgruppe von B enthält. Nach (8D) ist der Defekt von B mindestens gleich dem Defekt d von $\tilde{B}$. Da aber $\mathfrak{D}$ von der Ordnung p^d eine Sylowgruppe von $\mathfrak{N}(H)$ ist, so muß B den Defekt d haben und $\mathfrak{D}$ Defektgruppe von B sein.

Nehmen wir jetzt an, daß es noch einen zweiten Block $\tilde{B}_1$ von $\mathfrak{H}$ vom Defekt d gäbe, dem derselbe Block B von $\mathfrak{G}$ zugeordnet ist; es sei $\tilde{\psi}_1$ der zu $\tilde{B}_1$ gehörige lineare Charakter von $Z(\mathfrak{H}, \Omega^*)$. Dann hat auch $\tilde{B}_1$ die Defektgruppe $\mathfrak{D}$. Es sei $\mathfrak{L}_j$ eine Klasse von $\mathfrak{H}$ vom Defekt d in $\mathfrak{H}$. Gibt es in $\mathfrak{L}_j$ kein Element H, für das $\mathfrak{D}$ Sylowgruppe des Normalisators $\mathfrak{N}(H) \cap \mathfrak{H}$ von H in $\mathfrak{H}$ ist, so folgt aus (8A) daß $\tilde{\psi}(L_j) = 0$ und ebenso, daß $\tilde{\psi}_1(L_j) = 0$ ist. Gibt es in $\mathfrak{L}_j$ Elemente H, für die $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(H) \cap \mathfrak{H}$ ist, so gilt (10.1) und ebenso $\psi(K_i) = \tilde{\psi}_1(L_j)$, da ja B ebenfalls dem Block $\tilde{B}_1$ von $\mathfrak{H}$ zugeordnet ist; beim Beweis von (10.1) wurde die Bedingung $\tilde{\psi}(L_j) \neq 0$ nicht verwendet. Damit haben wir gezeigt, daß für alle Klassen $\mathfrak{L}_j$ von $\mathfrak{H}$ vom Defekt d die Gleichung

$$\tilde{\psi}(L_j) = \tilde{\psi}_1(L_j)$$

gilt. Dann folgt aber durch Anwendung von (6F) auf $\mathfrak{H}$, daß $\tilde{\psi} = \tilde{\psi}_1$, also $\tilde{B} = \tilde{B}_1$ ist, und damit ist (10B) in allen Teilen bewiesen.

Im Fall $d = 0$ ist unser Resultat (10B) inhaltslos, da hier $\mathfrak{D} = \{1\}$, $\mathfrak{T} = \mathfrak{H} = \mathfrak{G}$ ist. Es sei $d > 0$. Will man alle Blöcke vom Defekt d finden, so hat man zunächst ein volles System von nicht-konjugierten Untergruppen $\mathfrak{D}$ von $\mathfrak{G}$ von der Ordnung p^d aufzustellen. Selbstverständlich kann man alle $\mathfrak{D}$ in einer festen Sylowgruppe von $\mathfrak{G}$ wählen. Für jedes $\mathfrak{D}$ muß man dann $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$ bilden und alle Blöcke $\tilde{B}$ von $\mathfrak{H}$ vom Defekt d finden. Ist $\mathfrak{D}$ nicht invariant in $\mathfrak{G}$, so handelt es sich bei $\mathfrak{H}$ um eine Gruppe von kleinerer Ordnung als g. Wir werden aber in § 12 sehen, daß man die Behandlung von $\mathfrak{H}$ stets auf die von $\mathfrak{H}/\mathfrak{D}$ zurückführen kann, das jedenfalls kleinere Ordnung als $\mathfrak{G}$ hat. Damit wird dann gezeigt sein, daß die Aufstellung aller Blöcke mit gegebener Defektgruppe $\mathfrak{D} \neq \{1\}$ auf eine Untersuchung von Gruppen kleinerer Ordnung zurückgeführt werden kann.

Nicht jede Untergruppe der Ordnung p^d kann als Defektgruppe auftreten. Wir zeigen:

(10C) *Ist die Untergruppe $\mathfrak{D}$ von der Ordnung p^d Defektgruppe, so gibt es p-reguläre Elemente G von $\mathfrak{G}$, für die $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(G)$ ist. Ferner kann $\mathfrak{D}$ nicht in einer invarianten p-Untergruppe $\mathfrak{D}_1 \neq \mathfrak{D}$ von $\mathfrak{N}(\mathfrak{D})$ enthalten sein.*

Ist $\mathfrak{D}$ Defektgruppe des Blocks B, und verwenden wir dieselben Bezeichnungen wie früher, so zeigt (6B) daß es p-reguläre Klassen $\mathfrak{K}_\alpha$ vom Defekt d gibt, für die $\psi(K_\alpha) \neq 0$ ist. Dann ist $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(G)$ für geeignete

$G \in \mathfrak{K}_\alpha$. Gibt es eine invariante p-Untergruppe $\mathfrak{D}_1 > \mathfrak{D}$ von $\mathfrak{H}$, so enthält jede Defektgruppe eines Blocks $\tilde{B}$ von $\mathfrak{H}$ nach (9F) die Gruppe $\mathfrak{D}_1$, und daher besitzt $\mathfrak{H}$ keine Blöcke vom Defekt d.

§ 11. Fortsetzung von § 9

Nehmen wir an, daß die Gruppe $\mathfrak{G}$ eine invariante Untergruppe $\mathfrak{W}$ enthält. Ist ϑ ein Charakter von $\mathfrak{W}$ und G ein festes Element von $\mathfrak{G}$, so stellt für variables W in $\mathfrak{W}$ die Funktion $\vartheta_1(W) = \vartheta(GWG^{-1})$ wieder einen Charakter ϑ_1 von $\mathfrak{W}$ dar. Mit ϑ ist auch ϑ_1 irreduzibel. Wir sagen, daß die Charaktere ϑ und ϑ_1 von $\mathfrak{W}$ in $\mathfrak{G}$ *assoziiert* sind. Offenbar ist dies eine Äquivalenzrelation. Die Elemente G von $\mathfrak{G}$, für die $\vartheta(W) = \vartheta(GWG^{-1})$ für alle W in $\mathfrak{W}$ gilt, bilden eine Untergruppe $\mathfrak{S} \geqq \mathfrak{W}$ von $\mathfrak{G}$, die *Trägheitsgruppe* von ϑ in $\mathfrak{G}$. Die Anzahl r der verschiedenen zu ϑ in $\mathfrak{G}$ assoziierten Charaktere ist gleich dem Index $(\mathfrak{G}:\mathfrak{S})$.

Ist χ_i ein irreduzibler Charakter von $\mathfrak{G}$, so hat die Beschränkung $\chi_i^{(0)}$ von χ_i auf $\mathfrak{W}$ die Form

$$\chi_i^{(0)} = q \sum_{i=1}^{r} \vartheta_i, \tag{11.1}$$

wo $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ eine Klasse von in $\mathfrak{G}$ assoziierten Charakteren von $\mathfrak{W}$ bilden, und wo q eine natürliche Zahl ist. Hat ϑ den Grad t, so ergibt sich durch Gradvergleich die Gleichung

$$x_i = qrt. \tag{11.2}$$

Wir nehmen jetzt an, daß $\mathfrak{G}$ eine invariante Untergruppe $\mathfrak{D}$ der Ordnung p^d besitzt. Wir wollen die Blöcke B von $\mathfrak{G}$ untersuchen, die die Defektgruppe $\mathfrak{D}$ haben. Setzt man

$$\mathfrak{T} = \mathfrak{C}(\mathfrak{D}), \quad \mathfrak{W} = \mathfrak{T}\mathfrak{D}, \tag{11.3}$$

so sind auch $\mathfrak{T}$ und $\mathfrak{W}$ invariante Untergruppen von $\mathfrak{G}$.

Wie aus (9E) folgt, ist in B ein Block $\overline{B}$ von $\overline{\mathfrak{G}} = \mathfrak{G}/\mathfrak{D}$ vom Defekt $d - d = 0$ enthalten. Es sei χ_i der gewöhnliche irreduzible Charakter von $\overline{\mathfrak{G}}$ in B; wir betrachten wieder χ_i auch als Charakter von $\mathfrak{G}$. Ist $\nu(g) = a$, so teilt p die Ordnung von $\overline{\mathfrak{G}}$ mit dem Exponenten $a - d$, und daher hat man

$$\nu(x_i) = a - d. \tag{11.4}$$

Wir verwenden dieselben Bezeichnungen wie früher. Nach Definition der Defektgruppe gibt es Elemente G, für die $\mathfrak{D}$ Sylowgruppe von $\mathfrak{N}(G)$ ist, derart daß $\psi(K_\alpha) \not\equiv 0$ für die G enthaltende Klasse $\mathfrak{K}_\alpha$ von $\mathfrak{G}$ ist. Dann ist $\nu(G) = d$, und wir haben $G \in \mathfrak{C}(\mathfrak{D}) = \mathfrak{T}$. Da $\psi = \omega_i^*$ ist, so folgt daß $\omega_i(G) \not\equiv 0 \pmod{\mathfrak{p}}$ ist. Aus (2.3) folgt weiter wegen (11.4), daß $\chi_i(G) \not\equiv 0 \pmod{\mathfrak{p}}$ ist. Wendet man jetzt (11.1) an, so erhält man zunächst

$$q \not\equiv 0 \pmod{p}. \tag{11.5}$$

Da $\mathfrak{D}$ im Kern von χ_i enthalten ist, so ist $\mathfrak{D}$ auch im Kern von jedem ϑ_ϱ enthalten, und deswegen kann ϑ_ϱ als irreduzible Darstellung von $\mathfrak{W}/\mathfrak{D}$ betrachtet werden. Daher teilt der Grad t von ϑ_ϱ die Ordnung von $\mathfrak{W}/\mathfrak{D}$. Hat $\mathfrak{W}$ die Ordnung w, so liefert dies $\nu(t) \leq \nu(w) - d$. Andererseits folgt aus (11.2) wegen (11.4) und (11.5), daß

$$a - d = \nu(x_i) = \nu(r) + \nu(t)$$

ist. Also muß $a \leq \nu(r) + \nu(w)$ sein. Da $r = (\mathfrak{G} : \mathfrak{S})$ war und $\mathfrak{G} \supseteq \mathfrak{S} \supseteq \mathfrak{W}$ ist, ist dies nur möglich, wenn wir

$$\nu\big((\mathfrak{S} : \mathfrak{W})\big) = 0; \quad \nu(t) = \nu(w) - d \tag{11.6}$$

haben. Aus der zweiten Gl. (11.6) folgt, daß jedes ϑ_ϱ als Charakter von $\mathfrak{W}/\mathfrak{D}$ zu einem Block vom Defekt 0 gehört.

Umgekehrt sei jetzt ϑ ein Charakter von $\mathfrak{W}/\mathfrak{D}$, der zu einem Block von $\mathfrak{W}/\mathfrak{D}$ vom Defekt 0 gehört. Es sei $\mathfrak{S}$ die Trägheitsgruppe von ϑ in $\mathfrak{G}$, und es werde vorausgesetzt, daß $(\mathfrak{S} : \mathfrak{W})$ zu p teilerfremd ist. Dann gilt also (11.6). Ist $X_1, X_2, \ldots, X_r$ ein rechtsseitiges Restsystem von $\mathfrak{G}$ nach $\mathfrak{S}$, und setzt man

$$\vartheta_i(W) = \vartheta(X_i W X_i^{-1}) \qquad (\text{für } W \in \mathfrak{W}), \tag{11.7}$$

so sind $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ die verschiedenen, zu ϑ in $\mathfrak{G}$ assoziierten Charaktere. Bildet man den von ϑ induzierten Charakter ζ von $\mathfrak{G}$, so ist

$$\zeta(W) = \frac{1}{w} \sum_{X \in \mathfrak{G}} \vartheta(X W X^{-1}) = (\mathfrak{S} : \mathfrak{W}) \sum_{\varrho=1}^{r} \vartheta_\varrho(W) \qquad (\text{für } W \in \mathfrak{W}).$$

Ist χ_i ein irreduzibler, in ζ auftretender Charakter von $\mathfrak{G}$, so müssen also in der Formel (11.1) für χ_i gerade dieselben Charaktere $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ wie hier erscheinen. Da $\mathfrak{D}$ im Kern von ϑ und daher im Kern von allen ϑ_ϱ enthalten ist, gehört es auch zum Kern von ζ und dann auch von χ_i. Also kann χ_i auch als Charakter von $\mathfrak{G}/\mathfrak{D}$ betrachtet werden. Daher ist der Grad x_i von χ_i ein Teiler von g/p^d, also $\nu(x_i) \leq a - d$. Aus (11.1) folgt (11.2), und daraus erhalten wir unter Verwendung von (11.6) die Beziehung

$$\begin{aligned} \nu(x_i) &= \nu(q) + \nu(r) + \nu(t) = \nu(q) + \nu\big((\mathfrak{G} : \mathfrak{S})\big) + \nu(t) \\ &= \nu(q) + \nu\big((\mathfrak{G} : \mathfrak{W})\big) + \nu(w) - d = \nu(q) + a - d. \end{aligned}$$

Daher muß $\nu(x_i) = a - d$, $\nu(q) = 0$ sein. Es gelten also wieder (11.4) und (11.5).

Es sei B der Block von $\mathfrak{G}$, zu dem χ_i gehört. Wir wollen zeigen, daß B den Defekt d hat. Dazu brauchen wir den Hilfssatz:

(11A) *Es sei $\mathfrak{D}$ eine invariante Untergruppe der Ordnung p^d von $\mathfrak{G}$ und es sei $\mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, $\mathfrak{W} = \mathfrak{T}\mathfrak{D}$. Ist $\overline{T} = T\mathfrak{D}$ ein p-reguläres Element von $\mathfrak{W}/\mathfrak{D}$, so kann man den Repräsentanten T in $T\mathfrak{D}$ als p-reguläres Element wählen. Ist dies geschehen, so ist der Normalisator $\mathfrak{N}(\overline{T})$ von $\overline{T}$ in $\mathfrak{G}$ gerade $\mathfrak{N}(T)/\mathfrak{D}$.*

Beweis. Ist $\overline{T}$ p-regulär, so gibt es eine zu p teilerfremde, natürliche Zahl c, für die $\overline{T}^c = 1$ ist. Ist T ein Element der Restklasse $\overline{T}$, so ist $T^c \in \mathfrak{D}$, und da c zur Ordnung von $\mathfrak{D}$ teilerfremd ist, kann man eine Potenz

T_0 von T^e finden, für die $T_0^e = T^e$ ist. Da T und T_0 vertauschbar sind, ist $T_1 = TT_0^{-1}$ p-regulär. Offenbar ist $T_1\mathfrak{D} = T\mathfrak{D}$ und damit ist der erste Teil von (11A) bewiesen.

Ist T schon selbst p-regulär, und gehört $\overline{V} = V\mathfrak{D}$ zum Normalisator von $\overline{T}$ in $\overline{\mathfrak{G}}$, so ist $V^{-1}TV \in T\mathfrak{D}$; man setze $V^{-1}TV = TD$ mit $D \in \mathfrak{D}$. Hier ist das Element auf der linken Seite p-regulär. Auf der rechten Seite ist T p-regulär, während die Ordnung von D eine Potenz von p ist. Da D und T vertauschbar sind, kann TD nur dann p-regulär sein, wenn $D = 1$, d.h. wenn $V \in \mathfrak{N}(T)$ ist. Also ist $\mathfrak{N}(\overline{T}) \subseteq \mathfrak{N}(T)/\mathfrak{D}$. Da $\mathfrak{N}(T)/\mathfrak{D}$ sicher zu $\mathfrak{N}(\overline{T})$ gehört, ist auch der zweite Teil von (11A) bewiesen.

Nach Voraussetzung ist ϑ ein Charakter vom Defekt 0 von $\mathfrak{W}/\mathfrak{D}$. Wendet man den inneren Automorphismus $G \to XGX^{-1}$ von $\mathfrak{G}$ an, der $\mathfrak{W}$ in sich und ϑ in ϑ_i überführt, so sieht man, daß das gleiche für die r Charaktere ϑ_i gilt. Wie in (6G) gezeigt wurde, gibt es dann r p-reguläre Elemente $\overline{T}_1, \overline{T}_2, \ldots, \overline{T}_r$ von $\mathfrak{W}/\mathfrak{D}$, derart daß

$$\text{Det}\left(\vartheta_i(\overline{T}_j)\right) \not\equiv 0 \pmod{\mathfrak{p}}, \qquad (i, j = 1, 2, \ldots, r) \tag{11.8}$$

ist, und daß die Ordnung von $\mathfrak{N}(\overline{T}_j)$ zu p teilerfremd ist. Dann zeigt (11A), daß wir T_j in der Restklasse $\overline{T}_j$ so wählen können, daß

$$\nu(T_j) = d \tag{11.9}$$

ist.

Nach (2.3) finden wir dann unter Verwendung von (11.1) und (11.2), daß

$$\omega_i(W) = \frac{g}{n(W)tr} \sum_{j=1}^{r} \vartheta_i(W) \tag{11.10}$$

für $W \in \mathfrak{W}$ ist. Für $W = T_j$ haben wir $\nu\left(g/(n(W)tr)\right) = \nu(g/x_i) - \nu(T_j) = 0$. Aus (11.8) folgt, daß wir nicht $\omega_i(T_j) \equiv 0 \pmod{\mathfrak{p}}$ für alle T_j haben können. Auf Grund von (6B) ergibt sich also aus (11.9), daß der Block B von $\mathfrak{G}$ höchstens den Defekt d haben kann. Nach (9F) ist der Defekt aber nicht kleiner als d, und daher muß er genau gleich d sein, wie behauptet war. Die Defektgruppe von B ist folglich $\mathfrak{D}$.

Wie (6F) zeigt, ist der Block eindeutig durch die Werte $\omega_i^*(G)$ für p-reguläre Elemente G vom Defekt d bestimmt. Dabei darf man sich auf Elemente G beschränken, die zu Klassen $\mathfrak{K}_\alpha$ gehören, die Elemente T von $\mathfrak{T}$ enthalten. Für die anderen Klassen ergibt sich nämlich durch Anwendung von (10B) und (7.3) mit $\mathfrak{H} = \mathfrak{G}$, daß ω_i^* verschwindet. Jetzt zeigt (11.10), daß B nur von $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ abhängt.

Schließlich sei ϑ' ein Charakter von $\mathfrak{W}/\mathfrak{D}$, der zu einem Block vom Defekt 0 gehört, aber von $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ verschieden ist. Hat dann $\chi_{i'}$ die entsprechende Bedeutung für ϑ' wie χ_i für ϑ, und gehört $\chi_{i'}$ zum Block B' von $\mathfrak{G}$, so behaupten wir, daß B von B' verschieden ist. Um dies zu zeigen, hat man (6G) zur Bildung einer Determinante zu verwenden, in der die verschiedenen

zu ϑ und ϑ' assoziierten Charaktere von $\mathfrak{W}$ auftreten. Vergleicht man dann (11.10) mit der entsprechenden Formel für $\omega_{i'}$, so sieht man, daß ω_i^* für geeignete Elemente T von ω_i^* verschieden ist, und dann muß $B \neq B'$ sein.

Hat der Charakter ϑ von $\mathfrak{W}/\mathfrak{D}$ als Charakter von $\mathfrak{W}$ die Trägheitsgruppe $\mathfrak{S}$ in $\mathfrak{G}$, so hat er als Charakter von $\mathfrak{W}/\mathfrak{D}$ die Trägheitsgruppe $\mathfrak{S}/\mathfrak{D}$ in $\mathfrak{G}/\mathfrak{D}$. Damit erhalten wir den Satz:

(11B) *Es sei $\mathfrak{G}$ eine Gruppe, die eine invariante Untergruppe $\mathfrak{D}$ der Ordnung p^d besitzt. Es sei $\mathfrak{T}=\mathfrak{C}(\mathfrak{D})$, $\mathfrak{W}=\mathfrak{T}\mathfrak{D}$. Um alle Blöcke von $\mathfrak{G}$ vom Defekt d zu erhalten, hat man alle Charaktere ϑ vom Defekt 0 von $\mathfrak{W}/\mathfrak{D}$ zu finden, für die $\mathfrak{W}/\mathfrak{D}$ in der Trägheitsgruppe $\mathfrak{S}/\mathfrak{D}$ von ϑ einen zu p teilerfremden Index hat. Verteilen sich diese Charaktere auf u Familien von in $\mathfrak{G}/\mathfrak{D}$ assoziierten Charakteren, so gibt es u Blöcke B von $\mathfrak{G}$ vom Defekt d. Ist B der durch die Familie $\vartheta_1, \vartheta_2, \ldots, \vartheta_r$ von assoziierten Charakteren vom Grad t von $\mathfrak{W}/\mathfrak{D}$ bestimmte Block, und ist χ_i ein Charakter von $\mathfrak{G}$, so ist für Elemente $W \in \mathfrak{W}$:*

$$\chi_i(W) = \frac{x_i}{gt} \sum_{X \in \mathfrak{G}} \vartheta(XWX^{-1}). \tag{11.11}$$

Dabei erhält man (11.11) aus (11.1), indem man X über alle g Elemente von $\mathfrak{G}$ an Stelle über das Restsystem X_i laufen läßt und dementsprechend durch g/r dividiert; $qr/g = x_i/(gt)$.

§ 12. Kongruenzen für die Werte der Charaktere

Es sei wieder $\mathfrak{G}$ eine beliebige endliche Gruppe, und es sei B ein Block von $\mathfrak{G}$ vom Defekt d mit der Defektgruppe $\mathfrak{D}$. Setzt man $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$, $\mathfrak{T}=\mathfrak{C}(\mathfrak{D})$, $\mathfrak{W}=\mathfrak{T}\mathfrak{D}$, so gibt es nach (10B) einen Block $\tilde{B}$ von $\mathfrak{H}$ vom Defekt d, dem B zugeordnet ist. Ist χ_i ein Charakter von B und $\tilde{\chi}_i$ ein Charakter von $\tilde{B}$, so ist nach (7.4)

$$\chi_i(G) \equiv \frac{n(G)}{g} \frac{x_i}{\tilde{x}_j} \sum_T \tilde{\chi}_j(T) \qquad (\mathrm{mod}\, \mathfrak{p}_\nu \mathfrak{p}^{\nu(G)-\nu(g/x_i)}), \tag{12.1}$$

wo T über alle Elemente von $\mathfrak{T}$ läuft, die in $\mathfrak{G}$ zu G konjugiert sind. Da $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$ die invariante Untergruppe $\mathfrak{D}$ hat, kann man auf $\mathfrak{H}$ das Resultat (11B) anwenden. Ist also der Block $\tilde{B}$ von $\mathfrak{H}$ durch den Charakter ϑ von $\mathfrak{W}/\mathfrak{D}$ vom Defekt 0 bestimmt, so hat man für $\tilde{\chi}_j$ an Stelle von χ_i eine Formel (11.11):

$$\tilde{\chi}_j(T) = \frac{\tilde{x}_j}{ht} \sum_{X \in \mathfrak{H}} \vartheta(XTX^{-1}),$$

wo h die Ordnung von $\mathfrak{H}$ ist. Setzt man dies in (12.1) ein, so hat man

$$\chi_i(G) \equiv \frac{n(G)\, x_i}{ght} \sum_T \sum_{X \in \mathfrak{H}} \vartheta(XTX^{-1}) \qquad (\mathrm{mod}\, \mathfrak{p}_\nu \mathfrak{p}^{\nu(G)-\nu(g/x_i)}).$$

Jedes Glied in der Summe auf der rechten Seite hat die Form $\vartheta(T')$ mit $T'=XTX^{-1}$; $T \in \mathfrak{T}$ zu G in $\mathfrak{G}$ konjugiert; $X \in \mathfrak{H}$. Daher ist auch T' zu G

konjugiert und $T' \in \mathfrak{T}$. Ein festes T' wird genau h-mal in der Form XTX^{-1} erhalten. Schreibt man also für T' wieder T, so haben wir

$$\chi_i(G) \equiv \frac{n(G)}{g} \frac{x_i}{t} \sum_T \vartheta(T) \qquad (\mathrm{mod}\, \mathfrak{p}, p^{\nu(G)-\nu(g/x_i)}), \tag{12.2}$$

wo T dieselben Werte wie in (12.1) durchläuft.

Man kann dies Resultat noch in etwas anderer Form schreiben. Der von ϑ (als Charakter von $\mathfrak{W}$) in $\mathfrak{G}$ induzierte Charakter Θ ist durch die Formel

$$w\Theta(G) = \sum_{X \in \mathfrak{G}} \vartheta(XGX^{-1})$$

definiert, wo X über die Elemente von $\mathfrak{G}$ läuft, für die $XGX^{-1} \in \mathfrak{W}$ ist, und wo w die Ordnung von $\mathfrak{W}$ ist. Man betrachte hier zunächst die Summe Σ_1 der Glieder, für die XGX^{-1} ein Element von $\mathfrak{T}$ ist. Da jedes feste T $n(G)$-mal auftritt, ist $(1/n(G))\Sigma_1$ gerade die in (12.2) auftretende Summe.

Es sei weiter Σ_2 die Summe der übrigbleibenden Glieder,

$$\Sigma_2 = \Sigma\vartheta(XGX^{-1}); \qquad XGX^{-1} \in \mathfrak{W}, \qquad XGX^{-1} \notin \mathfrak{T}.$$

Hier setze man $XGX^{-1} = W = TD_0$ mit $T \in \mathfrak{T}$, $D_0 \in \mathfrak{D}$ und schreibe $T = AB$, wo A und B Potenzen von T sind, derart daß die Ordnung von A zu p teilerfremd ist und die von B eine Potenz von p. Dann liegen A und B in $\mathfrak{T}$, sind also mit D_0 vertauschbar. Deutet weiterhin ein Querstrich Bildung der Restklasse nach $\mathfrak{D}$ an, so haben wir

$$W = ABD_0, \qquad \overline{W} = \overline{AB}.$$

Ist nun $B \notin \mathfrak{D}$, also $\overline{B} \neq \overline{1}$, so ist $\overline{W}$ nicht p-regulär. Da ϑ ein Charakter von $\mathfrak{W}/\mathfrak{D}$ vom Defekt 0 war, ist $\vartheta(\overline{W}) = \vartheta(W) = 0$, vgl. (6E). Daher darf man sich in Σ_2 auf diejenigen Glieder beschränken, für die $XGX^{-1} = W$ die Form AD mit $D \in \mathfrak{D}$ und einem p-regulären $A \in \mathfrak{T}$ hat. Wegen $W \notin \mathfrak{T}$ muß $D \neq 1$ sein; wir haben $\vartheta(W) = \vartheta(A)$. Jedes derartige W tritt $n(G)$-mal auf. Mit jedem W fasse man die Glieder W' zusammen, die in $\mathfrak{W}$ zu W konjugiert sind. Diese erfüllen dieselben Bedingungen wie W; $W' = A'D'$ mit $D' \in \mathfrak{D}$ und p-regulären $A' \in \mathfrak{T}$, $D' \neq 1$. Ferner ist $\vartheta(W) = \vartheta(W')$. Zu jedem W gibt es $w/n_0(W)$ verschiedene W', wo $n_0(W)$ die Ordnung des Normalisators $\mathfrak{N}_0(W) = \mathfrak{N}(W) \cap \mathfrak{W}$ von W in $\mathfrak{W}$ ist. Da A und D vertauschbar sind und teilerfremde Ordnungen haben, ist A eine Potenz von W und also $\mathfrak{N}_0(W) \subseteq \mathfrak{N}_0(A) = \mathfrak{N}(A) \cap \mathfrak{W}$. Wir behaupten, daß $\nu(n_0(W)) < \nu(n_0(A))$ ist, wo $n_0(A)$ die Ordnung von $\mathfrak{N}_0(A)$ bezeichnet.

In der Tat, würde Gleichheit bestehen, so würde eine Sylowgruppe $\mathfrak{P}_0$ von $\mathfrak{N}_0(W)$ gleichzeitig eine Sylowgruppe von $\mathfrak{N}_0(A)$ sein. Wegen $A \in \mathfrak{T} = \mathfrak{C}(\mathfrak{D})$ gibt es Sylowgruppen von $\mathfrak{N}_0(A)$, die $\mathfrak{D}$ enthalten. Da $\mathfrak{D}$ invariant in $\mathfrak{W}$ ist, enthalten also alle Sylowgruppen von $\mathfrak{N}_0(A)$ die Gruppe $\mathfrak{D}$. Man hätte also $\mathfrak{D} \subseteq \mathfrak{P}_0 \subseteq \mathfrak{N}_0(W)$, was $W \in \mathfrak{C}(\mathfrak{D}) = \mathfrak{T}$ zur Folge hätte; wir hatten aber $W \notin \mathfrak{T}$ in Σ_2. Also ist $\nu(n_0(W)) < \nu(n_0(A))$, wie behauptet war.

Schließlich ist $w\,\vartheta(A)/(n_0(A)\,t)$ ganz algebraisch, also

$$\nu(\vartheta(A)) + \nu(w) \geq \nu(n_0(A)) + \nu(t) > \nu(n_0(W)) + \nu(t).$$

Faßt man all dies zusammen, so ergibt sich, daß für die Summe Σ_2' der zu einem gegebenen W gehörigen Glieder $\vartheta(W')$ die Beziehung gilt

$$\nu(\Sigma_2') \geq \nu(G) + \nu(w) - \nu(n_0(W)) + \nu(\vartheta(A)) > \nu(G) + \nu(t).$$

Da Σ_2 eine Summe von Teilsummen Σ_2' ist, ist

$$\nu(\Sigma_2) > \nu(G) + \nu(t).$$

Wir können nun (12.2) in der Form schreiben

$$\chi_i(G) \equiv \frac{1}{g}\frac{x_i}{t}\,(w\,\Theta(G) - \Sigma_2) \pmod{\mathfrak{p},\, p^{\nu(G)-\nu(g/x_i)}}.$$

Jetzt sieht man, daß man hier das Glied Σ_2 weglassen darf. Wir haben damit das folgende Resultat bewiesen

(12A) *Es sei $\mathfrak{D}$ eine Untergruppe von $\mathfrak{G}$, deren Ordnung eine Potenz p^d der Primzahl p ist; es sei $\mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$. Man bestimme ein volles System von Charakteren $\vartheta_1, \vartheta_2, \ldots$ von $\mathfrak{T}\mathfrak{D}/\mathfrak{D}$ vom Defekt 0, derartig daß ϑ_i in $\mathfrak{H}/\mathfrak{D}$ eine Trägheitsgruppe $\mathfrak{S}_i/\mathfrak{D}$ hat, für die der Index $(\mathfrak{S}_i/\mathfrak{T}\mathfrak{D})$ zu p teilerfremd ist. In diesem System finde man ein vollständiges Teilsystem $\vartheta^{(1)}, \vartheta^{(2)}, \ldots, \vartheta^{(u)}$ von Charakteren ϑ_i von denen keine zwei in $\mathfrak{H}/\mathfrak{D}$ assoziiert sind. Jedem $\vartheta^{(\varrho)}$ kann man dann einen Block $B^{(\varrho)}$ von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$ zuordnen, und alle Blöcke B von $\mathfrak{G}$ mit der Defektgruppe $\mathfrak{D}$ werden genau einmal in der Form $B^{(\varrho)}$ erhalten. Ist χ_i ein gewöhnlicher irreduzibler Charakter in $B^{(\varrho)}$, so ist*[16])

$$\chi_i(G) \equiv \frac{1}{(\mathfrak{G}:\mathfrak{T}\mathfrak{D})}\frac{x_i}{t^{(\varrho)}}\,\Theta^{(\varrho)}(G) \pmod{\mathfrak{p},\, p^{\nu(G)-\nu(g/x_i)}}. \tag{12.3}$$

Hier bezeichnet $\Theta^{(\varrho)}$ den von $\vartheta^{(\varrho)}$ induzierten Charakter von $\mathfrak{G}$, wobei $\vartheta^{(\varrho)}$ wieder als Charakter von $\mathfrak{T}\mathfrak{D}$ aufzufassen ist; x_i ist der Grad von χ_i und $t^{(\varrho)}$ der von $\vartheta^{(\varrho)}$.

§ 13. Über die Anzahl der Blöcke vom Defekt 0

Wie bereits bemerkt, versagen unsere Methoden im Fall der Blöcke vom Defekt 0. Daher entsteht die Aufgabe, die Bestimmung der Anzahl der Blöcke vom Defekt 0 zu diskutieren. Sind die Grade x_i der irreduziblen Charaktere χ_i bekannt, so kann man unmittelbar ablesen, welche von ihnen durch die volle Potenz p^a von p in g teilbar sind. Die Anzahl dieser x_i ist dann die gesuchte Anzahl M. Natürlich ist M vollständig durch die Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$ bestimmt. Das folgende Resultat (13A) drückt dies in einer

[16]) Die Bedeutung der Kongruenz (12.3) ist wieder die folgende: Ist δ die Differenz der linken und rechten Seite von (12.3), so ist

$$\nu(\delta) > \nu(G) - \nu(g/x_i).$$

konkreteren Form aus. Wir sind aber nicht in der Lage, die Bestimmung von M ähnlich wie im Fall eines positiven Defekts durchzuführen.

(13A) *Die Elemente K_α der Klassenalgebra $Z(\mathfrak{G}, \Omega^*)$, die zu Klassen $\mathfrak{K}_\alpha$ vom Defekt 0 gehören, bilden die Basis einer Teilalgebra U. Die Anzahl der Blöcke von $\mathfrak{G}$ vom Defekt 0 ist gleich der Dimension von U^n für genügend großes n.*

Beweis. Für jede p-Untergruppe $\mathfrak{D}$ von $\mathfrak{G}$ liefert Anwendung von (7C) mit $\mathfrak{T} = \mathfrak{C}(\mathfrak{D})$, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$ ein Ideal $\mathfrak{U}(\mathfrak{D})$ von $Z(\mathfrak{G}, \Omega^*)$, dessen Basis aus denjenigen K_α besteht, für die $\mathfrak{K}_\alpha \cap \mathfrak{T}$ leer ist. Offenbar ist dann U der Durchschnitt aller dieser $\mathfrak{U}(\mathfrak{D})$ mit $\mathfrak{D} \neq (1)$. Dies zeigt, daß U ein Ideal von $Z(\mathfrak{G}, \Omega^*)$ und daher jedenfalls eine Teilalgebra ist.

Jeder lineare Charakter ψ_0 von U kann zu einem linearen Charakter ψ von $Z(\mathfrak{G}, \Omega^*)$ erweitert werden. Gehörte hier ψ zu einem Block B von positivem Defekt, so würde nach (6B) ψ für alle K_α verschwinden, für die $\mathfrak{K}_\alpha$ Defekt 0 hat. Dann würde ψ für alle Elemente von U verschwinden, was unmöglich ist. Daher muß ψ zu einem Block vom Defekt 0 gehören.

Geht man umgekehrt von einem Block B vom Defekt 0 aus, so zeigt (6G), daß das zugehörige ψ nicht in U identisch verschwindet und ferner, daß zwei derartige ψ nicht dieselbe Beschränkung auf U haben können. Dies zeigt, daß die Anzahl M der Blöcke von $\mathfrak{G}$ vom Defekt 0 gleich der Anzahl der verschiedenen linearen Charaktere von U ist.

Es sei weiter ψ_0 ein linearer Charakter von U und ψ seine Erweiterung zu einem linearen Charakter von $Z(\mathfrak{G}, \Omega^*)$. Da ψ zu einem Block B von $\mathfrak{G}$ vom Defekt 0 gehört, so folgt aus (6E) und dem in § 4 Gesagten, daß ψ nur ein einziges Mal als irreduzibler Bestandteil in der regulären Darstellung $\mathfrak{R}$ von $Z(\mathfrak{G}, \Omega^*)$ auftritt. Ist $\mathfrak{R}_0$ die Beschränkung von $\mathfrak{R}$ auf U, so tritt dann ψ_0 nur ein einziges Mal in $\mathfrak{R}_0$ auf, da ψ eindeutig durch ψ_0 bestimmt war. Nun ist die reguläre Darstellung $\mathfrak{R}_U$ von U ein Bestandteil von $\mathfrak{R}_0$. Es folgt also, daß die reguläre Darstellung $\mathfrak{R}_U$ der kommutativen Algebra U jeden ihrer M linearen Charaktere ψ_0 mit der Vielfachheit 1 enthält. Dann zeigt aber die Strukturtheorie der Algebren, daß U direkte Summe des Radikals von U und von M Idealen ist, die als Algebren zu Ω^* isomorph sind. Daraus ergibt sich, daß U^n für genügend großes n in der Tat die Dimension M hat.

§ 14. Anhang: Bemerkungen über Algebren

Um die Arbeit möglichst unabhängig zu machen, geben wir hier einige Bemerkungen über Algebren, die wir aber nur so weit verfolgen, wie es in unserem Zusammenhang nötig ist. In § 14 bezeichnet Γ eine Algebra endlichen Ranges über einem Grundkörper Ω.

(a) Unter einem Ideal $\mathfrak{M}$ von Γ wird stets ein zweiseitiges Ideal verstanden. Ein Ideal $\mathfrak{M} \neq \Gamma$ ist ein *Primideal*, wenn ein Produkt $\mathfrak{U}\mathfrak{V}$ zweier Ideale $\mathfrak{U}$, $\mathfrak{V}$ von Γ nur dann in $\mathfrak{M}$ enthalten ist, falls mindestens einer der Faktoren in $\mathfrak{M}$ enthalten ist. Ist Γ halbeinfach und also direkte Summe $\Gamma = B_1 \oplus \cdots \oplus B_s$ von Idealen, die einfache Ringe sind, so sieht man sofort,

daß man genau s Primideale hat. Dies sind die Ideale $\mathfrak{P}_j = \sum_{\lambda \neq j} B_\lambda$. Ist Γ nicht halbeinfach und ist N das Radikal, so folgt aus der Nilpotenz des Radikals leicht, daß N in jedem Primideal von Γ enthalten ist. Ist $\Gamma \neq N$, hat also z.B. Γ ein Einheitselement, so sind die Primideale von Γ gerade diejenigen Ideale, die bei dem natürlichen Homomorphismus von Γ auf Γ/N den Primidealen von Γ/N entsprechen. Daraus ergibt sich, daß jedes Primideal $\mathfrak{P}$ von Γ maximal in Γ ist; die Umkehrung ist trivial. Ist $\mathfrak{P}$ ein Primideal von Γ, so ist $\Gamma/\mathfrak{P}$ eine einfache Algebra und hat daher genau eine irreduzible Darstellung $\mathfrak{F}$ (wobei ähnliche Darstellungen als nicht verschieden betrachtet sind). Dann kann $\mathfrak{F}$ als irreduzible Darstellung von Γ mit dem Kern $\mathfrak{P}$ betrachtet werden. Umgekehrt ist der Kern einer irreduziblen Darstellung $\mathfrak{F}$ von Γ notwendig ein maximales Ideal $\mathfrak{P}$ in Γ. Es besteht also eine eineindeutige Zuordnung zwischen den irreduziblen Darstellungen $\mathfrak{F}$ von Γ und den Primidealen $\mathfrak{P}$ von Γ. Dabei ist jeder irreduziblen Darstellung $\mathfrak{F}$ der Kern $\mathfrak{P}$ von $\mathfrak{F}$ zugeordnet.

(b) Hat eine von der Nulldarstellung verschiedene Darstellung φ von Γ den Grad 1, so sprechen wir auch von einem *linearen Charakter* von Γ. Diese linearen Charaktere können dann also auch als die Homomorphismen der Algebra Γ auf Ω definiert werden.

Aus dem Schurschen Lemma ergibt sich leicht, daß jede irreduzible Darstellung $\mathfrak{F}$ von Γ als Bestandteil in der regulären Darstellung $\mathfrak{R}$ von Γ auftritt. Ist Γ kommutativ, so folgt aus dem Schurschen Lemma weiter, daß nach einer geeigneten Grundkörpererweiterung alle irreduziblen Darstellungen von Γ den Grad 1 haben. Gilt dies bereits für den Grundkörper Ω, so sagen wir, daß Ω *Zerfällungskörper* von Γ ist. Dann zerfällt also die reguläre Darstellung $\mathfrak{R}$ von Γ in Bestandteile ersten Grades (nicht notwendig vollständig).

(14A) *Ist Γ eine kommutative Algebra, für die der Grundkörper Zerfällungskörper ist, so bildet jeder von Null verschiedene Homomorphismus ψ von Γ in einen Erweiterungskörper von Ω notwendig Γ auf Ω ab; es ist also ψ ein linearer Charakter von Γ. Ist B eine Teilalgebra von Γ, so kann jeder lineare Charakter von B zu einem linearen Charakter von Γ erweitert werden.*

Beweis. Bildet ψ die Algebra Γ auf den Körper $\Omega_1 \supseteq \Omega$ ab, so hat Ω_1 notwendig endlichen Grad m über Ω. Da Ω_1 als Algebra über Ω eine irreduzible Darstellung m-ten Grades hat, so kann man ψ selbst als irreduzible Darstellung m-ten Grades von Γ deuten. Nach Voraussetzung ist Ω Zerfällungskörper von Γ und also ist $m = 1$, d.h. $\Omega_1 = \Omega$.

Es sei $\mathfrak{R}^0$ die durch Beschränkung der regulären Darstellung $\mathfrak{R}$ von Γ auf die Teilalgebra B entstehende Darstellung von B. Es zerfällt also $\mathfrak{R}^0$ in Bestandteile ersten Grades. Diese Bestandteile sind die durch Beschränkung der Darstellungen ersten Grades von Γ entstehenden Darstellungen von B.

Ist $\tilde{\mathfrak{R}}$ die reguläre Darstellung von B, so sieht man sofort, daß $\tilde{\mathfrak{R}}$ als Bestandteil von $\mathfrak{R}^0$ aufgefaßt werden kann. Jeder lineare Charakter von B

ist ein Bestandteil von $\tilde{\mathfrak{R}}$, also von $\mathfrak{R}^0$, und ist daher die Beschränkung eines linearen Charakters φ von Γ auf B.

Damit ist (14A) bewiesen. Natürlich kann man auch einen direkten Beweis geben.

(c) Es sei Γ eine Algebra mit Einheitselement; wir fassen Ω als Teilkörper des Zentrums von Γ auf. In Ω sei ein fester Teilring $\mathfrak{i}$ gewählt, der das Einheitselement 1 enthält und Ω als Quotientenkörper besitzt. Unter einer *Ordnung* von Γ verstehen wir eine Teilmenge J, die die folgenden Bedingungen erfüllt:

(I) J ist ein Teilring von Γ; $1 \in J$.

(II) J ist ein $\mathfrak{i}$-Modul, der endlich erzeugt werden kann.

(III) J ist nicht in einer eigentlichen Teilalgebra von Γ enthalten.

Sprechen wir von einem Ideal von J, so meinen wir stillschweigend ein zweiseitiges Ideal, das von Null verschiedene Elemente von $\mathfrak{i}$ enthält.

Es sei $\mathfrak{T}$ eine Darstellung der Algebra Γ durch lineare Transformationen eines Vektorraumes V über Ω. Wählt man eine Basis $e_1, e_2, \ldots, e_n$ von V, so wird jede lineare Transformation $\mathfrak{T}(\alpha)$, $(\alpha \in \Gamma)$, durch eine Matrix $[\mathfrak{T}(\alpha)]$ mit Koeffizienten in Ω gegeben. Man erhält so eine Darstellung $[\mathfrak{T}]$ von Γ durch Matrizen. Man sagt, daß $[\mathfrak{T}]$ ganzzahlig für J ist, wenn jedes $[\mathfrak{T}(\alpha)]$ mit $\alpha \in J$ Koeffizienten in $\mathfrak{i}$ hat. Offenbar ist dies dann und nur dann der Fall, wenn der von $e_1, e_2, \ldots, e_n$ erzeugte $\mathfrak{i}$-Modul U durch alle linearen Transformationen $\mathfrak{T}(\alpha)$ mit $\alpha \in J$ in sich abgebildet wird. Schreibt man $v\alpha$ für das Bild eines Vektors $v \in V$ für die lineare Transformation $\mathfrak{T}(\alpha)$, so ist U ein J-Modul.

(d) Wir setzen weiter voraus, daß $\mathfrak{i}$ ein Hauptidealring ist, was für unsere Zwecke ausreicht. Dann besitzt jede Ordnung J eine $\mathfrak{i}$-Basis. Ist $\mathfrak{T}$ eine Darstellung von Γ und ist $e_1^{(0)}, \ldots, e_n^{(0)}$ eine Ω-Basis des zugrunde liegenden Vektorraumes V, so sei U der von den $e_i^{(0)}$ erzeugte J-Modul. Offenbar kann dann U auch als $\mathfrak{i}$-Modul endlich erzeugt werden und besitzt daher auch eine $\mathfrak{i}$-Basis $e_1, e_2, \ldots, e_n$. Diese bildet dann auch eine Ω-Basis von V. Legt man diese Basis bei der Bildung von $[\mathfrak{T}]$ zugrunde, so wird also $[\mathfrak{T}]$ eine ganzzahlige Darstellung von J. In jeder Klasse ähnlicher Darstellungen von Γ gibt es also ganzzahlige Darstellungen von J.

Es sei $\mathfrak{p}$ ein Primideal von $\mathfrak{i}$; der Restklassenkörper $\mathfrak{i}/\mathfrak{p}$ werde mit Ω^* bezeichnet. Man sieht leicht, daß $\mathfrak{i} \cap \mathfrak{p}J = \mathfrak{p}$ ist und daß $J^* = J/\mathfrak{p}J$ eine Algebra über Ω^* ist.

(e) Ist $[\mathfrak{T}]$ eine ganzzahlige Darstellung von J, so führt die Restklassenabbildung von $\mathfrak{i} \pmod{\mathfrak{p}}$ diese in eine Darstellung $\mathfrak{T}^*$ von J^* in Ω^* über. Wir zeigen

(14B) *Es sei* $\mathfrak{i}$ *der Ring* $\mathfrak{o}_\nu$ *der für die diskrete Bewertung* ν *des Körpers* Ω *ganzen Zahlen, und es sei* $\mathfrak{p}$ *das zugehörige Primideal. Sind* $[\mathfrak{T}_1]$ *und* $[\mathfrak{T}_2]$ *zwei ähnliche ganzzahlige Darstellungen der Ordnung* J, *so haben die zugehörigen*

„modularen" Darstellungen $\mathfrak{T}_1^*$ und $\mathfrak{T}_2^*$ dieselben irreduziblen Bestandteile (natürlich als Darstellungen von J^* im Körper Ω^*).

Beweis. Es gehöre etwa $[\mathfrak{T}_1]$ zu dem J-Modul U_1 und $[\mathfrak{T}_2]$ zu dem J-Modul U_2, die beide demselben Vektorraum V angehören, und die beide den Ω-Modul V aufspannen. Es sei π eine zu $\mathfrak{p}$ gehörige Primzahl. Da man U_2 durch cU_2 mit c in Ω, $c \neq 0$ ersetzen kann, so dürfen wir annehmen, daß $U_2 \subseteq U_1$ ist. Ferner gibt es eine natürliche Zahl r, derart daß $\pi^r U_1 \subseteq U_2$ ist;

$$\pi^r U_1 \subseteq U_2 \subseteq U_1.$$

Der Darstellungsraum für $\mathfrak{T}_1^*$ kann als $U_1/\pi U_1$ und der für $\mathfrak{T}_2^*$ als $U_2/\pi U_2$ gewählt werden. Man setze

$$X_m = (U_2 \cap \pi^m U_1) + \pi U_2; \qquad Y_m = (U_1 \cap \pi^{1-m} U_2) + \pi U_1.$$

Dann ist

$$U_2 = X_0 \supseteq X_1 \supseteq \cdots \supseteq X_{r+1} = \pi U_2,$$
$$\pi U_1 = Y_0 \subseteq Y_1 \subseteq \cdots \subseteq Y_{r+1} = U_1.$$

Die Faktorgruppen X_m/X_{m+1} und Y_{m+1}/Y_m können als J^*-Moduln betrachtet werden. In diesem Sinne zerfällt die Darstellung $\mathfrak{T}_1^*$ in $U_1/\pi U_1$ in die Darstellungen von J^*, die zu den von (0) verschiedenen Moduln Y_{m+1}/Y_m gehören. Ebenso zerfällt $\mathfrak{T}_2^*$ in die Darstellungen, die zu den von (0) verschiedenen Moduln X_m/X_{m+1} gehören. Können wir also zeigen, daß die J-Moduln Y_{m+1}/Y_m und X_m/X_{m+1} isomorph sind,

$$X_m/X_{m+1} \simeq Y_{m+1}/Y_m, \tag{14.1}$$

so sind wir fertig.

Wendet man nun das Zassenhaussche Lemma[17]) auf die J-Gruppen $\mathfrak{u} = \pi U_2$, $\mathfrak{U} = U_2$, $\mathfrak{v} = \pi^{m+1} U_1$, $\mathfrak{V} = \pi^m U_1$, an, so ergibt sich

$$\frac{X_m}{X_{m+1}} = \frac{(U_2 \cap \pi^m U_1) + \pi U_2}{(U_2 \cap \pi^{m+1} U_1) + \pi U_2} \simeq \frac{(\pi^m U_1 \cap U_2) + \pi^{m+1} U_1}{(\pi^m U_1 \cap \pi U_2) + \pi^{m+1} U_1}.$$

Bildet man weiter $(\pi^m U_1 \cap U_2) + \pi^{m+1} U_1$ durch den J-Isomorphismus $\xi \to \pi^{-m} \xi$ ab, so erhält man als Bild $(U_1 \cap \pi^{-m} U_2) + \pi U_1 = Y_{m+1}$, während der Teilmodul $(\pi^m U_1 + \pi U_2) + \pi^{m+1} U_1$ in $(U_1 \cap \pi^{1-m} U_2) + \pi U_1 = Y_m$ übergeht. Daher ist der Quotient auf der rechten Seite in unserer Beziehung zu Y_{m+1}/Y_m isomorph. Jetzt sieht man, daß (14.1) zunächst als J-Isomorphismus und dann auch als J^*-Isomorphismus gilt, und damit ist (14B) bewiesen.

Man kann den Beweis von (14B) auch leicht führen, indem man entweder zusammengehörige Basen der J-Moduln U_1 und U_2 verwendet oder auch, indem man die Elementarteilertheorie auf eine $[\mathfrak{T}_1]$ in $[\mathfrak{T}_2]$ transformierende Matrix anwendet.

(f) Schließlich zeigen wir

(14C) *Es sei Γ eine kommutative halbeinfache Algebra über dem Körper Ω, der Zerfällungskörper von Γ sei. Der Ring $\mathfrak{i}$ sei der Ring der für die diskrete*

[17]) Zassenhaus, H.: Lehrbuch der Gruppentheorie, S. 51. Leipzig u. Berlin 1937.

Bewertung v von Ω ganzen Zahlen, und es sei $\mathfrak{p}$ das zugehörige Primideal von $\mathfrak{i}$. Ist J eine Ordnung von Γ, so ist $\Omega^* = \mathfrak{i}/\mathfrak{p}$ Zerfällungskörper für die Ω^*-Algebra $J^* = J/\mathfrak{p}J$ und jeder lineare Charakter von J^* geht aus einem linearen Charakter von Γ durch Restklassenbildung hervor.

Beweis. Wählt man eine $\mathfrak{i}$-Basis von J zur Bildung der reguläten Darstellung $\mathfrak{R}$ von Γ, so ist $[\mathfrak{R}]$ eine ganzzahlige Darstellung von J. Sind $\omega_1, \omega_2, \ldots, \omega_n$ die verschiedenen linearen Charaktere von Γ, so ist $[\mathfrak{R}]$ ähnlich zu der Darstellung $[\mathfrak{R}_1]$, die vollständig in $\omega_1, \omega_2, \ldots, \omega_n$ zerfällt. Da jede Darstellung $[\omega_i]$ ganzzahlig ist [z. B. nach dem in (d) Gezeigten], so gilt das gleiche für $[\mathfrak{R}_1]$. Jetzt zeigt (14B), daß $\mathfrak{R}_1^*$ die irreduziblen Bestandteile $\omega_1^*, \ldots, \omega_n^*$ besitzt, die alle Grad 1 haben. Offenbar ist $\mathfrak{R}^*$ die reguläre Darstellung von J^*. Da jeder lineare Charakter ψ von J^* in ihr als Bestandteil auftritt, stimmt ψ mit einem der ω_i^* überein. Damit ist (14C) vollständig gezeigt.

Dept. of Mathematics, Harvard University, Cambridge 38, Mass. (U.S.A.)

(Eingegangen am 24. August 1955)

RADEMACHER, H.
Math. Zeitschr. Bd. 63, S. 445—463 (1956)

Zur Theorie der Dedekindschen Summen

I. SCHUR zum Gedächtnis

Von

HANS RADEMACHER[1])

1. Die Dedekindschen Summen $s(h, k)$, die in der Theorie der Dedekindschen Funktion $\log \eta(\tau)$ auftreten, sind definiert als

$$s(h,k) = \sum_{\mu \bmod k} \left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) \qquad (h,k)=1, \quad k>0, \tag{1}$$

wo für reelles x

$$((x)) = \begin{cases} 0 \text{ für ganzes } x \\ x - [x] - \frac{1}{2} \text{ sonst} \end{cases} \tag{2}$$

gesetzt ist. Ihre bemerkenswerteste Eigenschaft wird durch die Reziprozitätsformel

$$s(h,k) + s(k,h) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h}{k} + \frac{k}{h} + \frac{1}{hk}\right) \tag{3}$$

ausgedrückt, die einerseits aus dem Verhalten von $\log \eta(\tau)$ unter Modulsubstitutionen abzulesen ist, für die aber auch eine Anzahl direkter Beweise bekannt sind. Man kann (2) mit Hilfe eines Euklidischen Algorithmus, da $s(h_1, k) = s(h_2, k)$, für $h_1 \equiv h_2 \pmod k$ ist, zweckmäßigerweise zur numerischen Berechnung von $s(h, k)$ für größere Werte von k, wo die Summendefinition unbequem wäre, benutzen.

Es treten in (2) und in einigen Verallgemeinerungen, die in verschiedener Richtung von MORDELL [*1*] und von mir [*2*] gegeben worden sind, jedoch mehrere $s(h, k)$ zugleich auf. Im folgenden will ich nun Eigenschaften studieren, die der einzelnen Summe $s(h, k)$ zukommen. Abschnitt I soll sich mit den Werten von $s(h, k)$ beschäftigen, Abschnitt II von einer Anwendung der Dedekindschen Summen auf das Problem der Transformationsklassen der Modulgruppe handeln, und Abschnitt III zieht quadratische Formen heran.

I

2. Hilfssatz 1. *Wenn h' durch*

$$hh' \equiv 1 \pmod k \tag{4}$$

bestimmt wird, ist

$$s(h', k) = s(h, k). \tag{4a}$$

Ferner ist

$$s(-h, k) = -s(h, k). \tag{4b}$$

[1]) John Simon Guggenheim Memorial Fellow 1954/55.

Beweis. Die zweite Aussage folgt sofort aus der Tatsache, daß $((x))$ eine ungerade Funktion von x ist.

Wenn ferner μ ein volles Restsystem modulo k durchläuft, so tut dies auch $h'\mu$. Wir haben also

$$s(h,k)=\sum_{\mu \bmod k}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right)=\sum_{\mu \bmod k}\left(\left(\frac{h'\mu}{k}\right)\right)\left(\left(\frac{hh'\mu}{k}\right)\right)=\sum_{\mu \bmod k}\left(\left(\frac{h'\mu}{k}\right)\right)\left(\left(\frac{\mu}{k}\right)\right)=s(h',k).$$

3. Da $((x))$ eine ungerade Funktion der Periode 1 ist, folgt sofort

$$\sum_{\lambda \bmod k}\left(\left(\frac{\lambda}{k}\right)\right)=0$$

und somit

$$s(h,k)=\sum_{\mu=1}^{k-1}\left(\frac{\mu}{k}-\frac{1}{2}\right)\left(\left(\frac{h\mu}{k}\right)\right)=\sum_{\mu=1}^{k-1}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right).$$

Explizit heißt dies

$$(5)\qquad \begin{cases} s(h,k)=\displaystyle\sum_{\mu=1}^{k-1}\frac{\mu}{k}\left(\frac{h\mu}{k}-\left[\frac{h\mu}{k}\right]-\frac{1}{2}\right) \\ \quad=\dfrac{h}{k^2}\displaystyle\sum_{1}^{k-1}\mu^2-\frac{1}{k}\sum_{1}^{k-1}\mu\left[\frac{h\mu}{k}\right]-\frac{1}{2k}\sum_{1}^{k-1}\mu \\ \quad=\dfrac{h(k-1)(2k-1)}{6k}-\dfrac{1}{k}g-\dfrac{k-1}{4}, \end{cases}$$

wo g eine ganze Zahl ist. Man sieht daher, daß der Nenner von $s(h,k)$ höchstens $6k$ sein kann. Da ferner in (5) der Nenner des einzelnen Summanden höchstens k^2 für gerades k und $2k^2$ für ungerades k beträgt, so ist der Nenner von $s(h,k)$ höchstens

für k gerade: $\quad (6k,k^2)=k(6,k)=2k\cdot(3,k),$

für k ungerade: $\quad (6k,2k^2)=2k(3,k).$

Damit haben wir erhalten den

Hilfssatz 2. *Der Nenner von $s(h,k)$ ist höchstens $2k\cdot(3,k)$.*

Dieser Wert des Nenners wird erreicht, z.B. ist $s(1,3)=\frac{1}{18}$. Der Nenner kann aber auch ein echter Teiler des angegebenen Wertes sein: $s(3,25)=\frac{2}{5}$.

Multiplikation von (3) mit $12hk$ ergibt

$$(6)\qquad 12hk\,s(h,k)+12hk\,s(k,h)=-3hk+h^2+k^2+1.$$

Im Hinblick auf Hilfssatz 2 lesen wir hieraus leicht ab den

Hilfssatz 3. *Es sei $\Theta=(3,k)$. Dann ist*

$$(7)\qquad 12hk\,s(h,k)\equiv h^2+1 \pmod{\Theta k}.$$

4. Wir beweisen nun den

Satz 1. *Der einzige ganzzahlige Wert, den $12s(h,k)$ annimmt, ist Null. Dies tritt dann und nur dann ein, wenn*

$$(8)\qquad h^2+1\equiv 0 \pmod{k}.$$

Beweis. Es sei zunächst (8) erfüllt. Dann ist, in der Bezeichnung Hilfssatz 1

$$h' = -h,$$

also

$$s(h,k) = s(-h,k) = -s(h,k),$$

woraus die zweite Hälfte des Satzes folgt.

Sei nun umgekehrt $12s(h,k)$ ganz. Dann ist nach (6) und Hilfssatz 2 die Kongruenz (8) erfüllt.

5. Satz 2. *Es ist* $s(h,k) < s(1,k)$ *für* $1 < h < \sqrt{k}$.

Beweis. Wir haben nach (5)

$$k \cdot s(1,k) = \sum_{\mu=1}^{k-1} \mu\left(\frac{\mu}{k} - \frac{1}{2}\right),$$

$$k \cdot s(h,k) = \sum_{\mu=1}^{k-1} \mu\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2}\right).$$

Es genügt also zu zeigen, daß

$$\sum_{\mu=1}^{k-1} \mu^2 > \sum_{\mu=1}^{k-1} \mu\left(h\mu - k\left[\frac{h\mu}{k}\right]\right)$$

ist. Hier sind die Zahlen

$$a_\mu = h\mu - k\left[\frac{h\mu}{k}\right]$$

eine Permutation der Zahlen $1, 2, \ldots, k-1$. Für alle möglichen Permutationen $\{a_\mu\}$ der Zahlen 1 bis $k-1$ erreicht aber die Summe

$$\sum \mu a_\mu$$

ihr Maximum für die Anordnung $a_\mu = \mu$. In der Tat, wenn nicht durchgängig $\mu = a_\mu$ ist, dann gibt es in der Anordnung $a_1, a_2, \ldots, a_{k-1}$ ein größtes $\mu = \mu_0$, so daß $a_{\mu_0} \neq \mu_0$, während $a_\mu = \mu$ für $\mu > \mu_0$ (sofern ein solches μ vorrätig ist). Dann ist jedoch sogar $a_{\mu_0} < \mu_0$, denn die $a_\mu > \mu_0$ sind besetzt durch $a_\mu = \mu$. Es sei dann $\mu_0 = a_{\mu_1}$ mit $\mu_1 < \mu_0$. Wir vertauschen dann a_{μ_0} und a_{μ_1}, ersetzen also die beiden Summanden $\mu_1 a_{\mu_1} + \mu_0 a_{\mu_0}$ durch $\mu_1 a_{\mu_0} + \mu_0 a_{\mu_1}$, was eine Vergrößerung bedeutet:

$$\mu_1 a_{\mu_0} + \mu_0 a_{\mu_1} - (\mu_1 a_{\mu_1} + \mu_0 a_{\mu_0}) = (\mu_0 - \mu_1)(a_{\mu_1} - a_{\mu_0}) = (\mu_0 - \mu_1)(\mu_0 - a_{\mu_0}) > 0.$$

Satz 3. *Es ist* $s(h,k) > 0$ *für* $0 < h < \sqrt{k-1}$.

Beweis. Nach (3) und nach Satz 2 haben wir

$$s(h,k) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{hk} + \frac{k}{h}\right) - s(k,h)$$

$$\geq -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{hk} + \frac{k}{h}\right) - s(1,h).$$

Nun folgt aus (3), da $s(h, 1) = 0$ ist,

$$s(1, h) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{h} + \frac{1}{h}\right),$$

so daß wir erhalten

$$s(h, k) \geq \frac{(k-1)(k-h^2-1)}{12hk},$$

woraus die Behauptung des Satzes folgt.

Da in dieser Überlegung das Ungleichheitszeichen nur bei der Ersetzung von $s(k, h)$ durch $s(1, h)$ auftritt, so haben wir das

Korollar. $s(h, k) = \dfrac{(k-1)\left(\frac{k-1}{h} - h\right)}{12k}$ *für* $k \equiv 1 \pmod h$.

5. Satz 4. *Es seien* $\frac{h_1}{k_1} < \frac{h_2}{k_2}$ *zwei benachbarte Brüche in einer Farey-Reihe. Dann ist*

$$s(h_1, k_1) - s(h_2, k_2) = \frac{1}{4} - \frac{1 + k_1^2 + k_2^2}{12 k_1 k_2}. \tag{9}$$

Beweis. Die kennzeichnende Eigenschaft für benachbarte Farey-Brüche $\frac{h_1}{k_1} < \frac{h_2}{k_2}$ ist

$$\begin{vmatrix} h_1 & h_2 \\ k_1 & k_2 \end{vmatrix} = -1,$$

so daß also

$$h_1 k_2 \equiv -1 \pmod{k_1},$$
$$h_2 k_1 \equiv 1 \pmod{k_2}.$$

Dann ist aber nach Hilfssatz 1

$$s(k_2, k_1) = -s(h_1, k_1),$$
$$s(k_1, k_2) = s(h_2, k_2),$$

womit (9) aus (3) folgt, wenn dort h, k durch k_1, k_2 ersetzt werden.

Satz 5. *Über* h_1/k_1 *und* h_2/k_2 *mögen die Voraussetzungen von Satz 4 gelten. Dann ist*

$$s(h_1 + h_2, k_1 + k_2) = \frac{1}{2}\{s(h_1, k_1) + s(h_2, k_2)\} + \frac{(k_2 - k_1)(1 + k_1^2 + k_1 k_2 + k_2^2)}{24 k_1 k_2 (k_1 + k_2)}. \tag{10}$$

Beweis. Die Anwendung von (9) auf $\frac{h_1}{k_1} < \frac{h_1+h_2}{k_1+k_2} < \frac{h_2}{k_2}$ ergibt

$$s(h_1, k_1) - s(h_1 + h_2, k_1 + k_2) = \frac{1}{4} - \frac{1 + k_1^2 + (k_1 + k_2)^2}{12 k_1 (k_1 + k_2)},$$

$$s(h_1 + h_2, k_1 + k_2) - s(h_2, k_2) = \frac{1}{4} - \frac{1 + (k_1 + k_2)^2 + k_2^2}{12 (k_1 + k_2) k_2}.$$

Durch Subtraktion dieser beiden Gleichungen voneinander ergibt sich (10).

Die Formel (10) eignet sich besonders gut zur schrittweisen tabellarischen Berechnung von $s(h, k)$, ausgehend von $s(0, 1) = s(1, 1) = 0$.

6. Die Sätze 4 und 5 legen nahe, $s(h/k)$ statt $s(h, k)$ zu schreiben. Dies kann man unbedenklich tun, wenn man vorübergehend die Bedingung $(h, k) = 1$ fallen läßt. In der Tat hängt dann $s(h, k)$ nur von dem Verhältnis $h:k$ ab, wie man ersieht aus

Hilfssatz 4. *Für ganzes positives q ist*

$$s(qh, qk) = s(h, k).$$

Dies gilt nur für die Definition (1) von $s(h, k)$, nicht etwa für Formel (5). Dieser Hilfssatz ist Theorem 1 in [3] und dort bewiesen. [Alle Formeln setzen aber bisher $(h, k) = 1$ voraus.] Damit ist eine Funktion $s(\varrho)$ für alle rationalen $\varrho = h/k$ definiert. Über den Verlauf dieser Funktion x macht der folgende Satz eine Aussage.

Satz 6. *Die für rationale $\varrho = h/k$ definierte Funktion $s(\varrho)$ ist in jedem Intervall nach oben und unten unbeschränkt.*

Beweis. In (9) möge h_1/k_1 festgehalten werden und h_2/k_2 eine Folge von benachbarten Brüchen durchlaufen die von oben gegen h_1/k_1 gehen (z. B. durch festgesetze Mediantenbildung). Mit k_1 fest, $k_2 \to +\infty$ geht die rechte Seite von (9) gegen $-\infty$, also $s(h_2/k_2) \to +\infty$. Ebenso, wenn man h_2/k_2 festhält und h_1/k_1 gegen h_2/k_2 gehen läßt, folgt $s(h_1/k_1) \to -\infty$.

II

7. Es sei Γ die Gruppe aller homogenen Modulsubstitutionen

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1.$$

Zwei Substitutionen M_1, M_2 heißen ähnlich, wenn ein $L \in \Gamma$ existiert, so daß

$$M_1 = L^{-1} M_2 L.$$

Die Gruppe Γ zerfällt durch die Ähnlichkeitsrelation in *Klassen*. Die Spur $a + d$ ist offenbar eine Invariante der Ähnlichkeitsklassen. In diesem Abschnitt behandeln wir eine weitere Invariante.

Satz 7. *Die Funktion*

$$(11)\quad \begin{cases} \Psi(M) = \Psi\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \qquad = \begin{cases} \dfrac{b}{d} & \text{für } c = 0, \\ \dfrac{a+d}{c} - 12 \operatorname{sign} c \cdot s(a, |c|) - 3 \operatorname{sign}(c(a+d)) & \text{für } c \neq 0 \end{cases} \end{cases}$$

ist eine Klasseninvariante. Außerdem ist

$$(12)\qquad \Psi(M) = \Psi(-M), \quad \Psi(M^{-1}) = -\Psi(M),$$

und Ψ ist eine ganze Zahl.

Beweis. Die Aussagen (12) folgen aus der Definition (11), wenn man noch

$$M^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

beachtet. Ferner kann man aus (7) entnehmen

$$12|c|\, s(a,|c|) \equiv a + d \pmod{c},$$

also

$$\frac{a+d}{|c|} - 12\, s(a,|c|)$$

ganzzahlig, woraus die Ganzzahligkeit von Ψ in (11) folgt.

Für die Invarianz von $\Psi(M)$ brauchen wir nur die Gleichungen

$$\Psi(S^{-1} M S) = \Psi(M), \tag{13}$$

$$\Psi(T^{-1} M T) = \Psi(M), \tag{14}$$

mit

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

zu beweisen.

Mit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

haben wir

$$S^{-1} M S = \begin{pmatrix} a-c & a-c+b-d \\ c & c+d \end{pmatrix}, \tag{13a}$$

$$T^{-1} M T = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{14a}$$

Dann ist (13) unmittelbar aus der Definition von Ψ abzulesen. Wir haben also nur noch (14) zu untersuchen. Zunächst bemerken wir, daß

$$12\, s(1,k) = -3 + k + \frac{2}{k}, \tag{15}$$

was aus dem Korollar zu Satz 3, oder auch direkt aus (3) folgt. Für den Beweis von (14) unterscheiden wir drei Fälle.

I. $c = 0$, dann ist $a = d = 1$. [$a = d = -1$ kann auf Grund von (12) beiseite gelassen werden.] Dann lautet (14), (14a):

$$\Psi\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \Psi\begin{pmatrix} 1 & 0 \\ -b & 1 \end{pmatrix} \tag{16}$$

oder, nach (11),

$$b = \begin{cases} 0 & \text{für } b = 0 \\ \dfrac{2}{-b} + 12 \operatorname{sign} b \cdot s(1,|b|) + 3 \operatorname{sign} b, \end{cases}$$

was in der Tat durch (15) für $k = |b|$ bestätigt wird.

II. $c \neq 0$, $b = 0$. Dann handelt es sich um

$$\Psi\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \overset{\cdot}{=} \Psi\begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix},$$

was aus (16) durch Vertauschung von c und $-b$ hervorgeht und somit unter I erledigt ist.

III. $b \neq 0$, $c \neq 0$. Hier ist also im Hinblick auf (11), (14), (14a) zu beweisen:

$$(17) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(a, |c|) + 12 \operatorname{sign} b \cdot s(d, |b|) \\ = \dfrac{a+d}{c} + \dfrac{a+d}{b} - 3 \operatorname{sign}\big(c(a+d)\big) - 3 \operatorname{sign}\big(b(a+d)\big). \end{cases}$$

Wenn $d = 0$ ist, dann muß $b = -c = \pm 1$ sein, und (17) ist dann trivialerweise erfüllt.

Wenn $d \neq 0$, dann kann wegen (12) $d > 0$ angenommen werden. Dann ergibt (3)

$$12\, s(d, |c|) + 12 s(|c|, d) = -3 + \frac{d}{|c|} + \frac{|c|}{d} + \frac{1}{d|c|},$$

$$12 s(d, |b|) + 12 s(|b|, d) = -3 + \frac{d}{|b|} + \frac{|b|}{d} + \frac{1}{d|b|}.$$

Die erste dieser Gleichungen multiplizieren wir mit $\operatorname{sign} c$, die zweite mit $\operatorname{sign} b$ und beachten, daß zufolge (4b)

$$s(h, k) = \operatorname{sign} h \cdot s(|h|, k)$$

ist, so daß

$$(18) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(d, |c|) + 12 s(c, d) = -3 \operatorname{sign} c + \dfrac{d}{c} + \dfrac{c}{d} + \dfrac{1}{dc}, \\ 12 \operatorname{sign} b \cdot s(d, |b|) + 12 s(b, d) = -3 \operatorname{sign} b + \dfrac{d}{b} + \dfrac{b}{d} + \dfrac{1}{db} \end{cases}$$

hervorgeht. Nach Hilfssatz 1 ist

$$s(a, |c|) = s(d, |c|),$$

$$s(c, d) = -s(b, d).$$

Addition der Gln. (18) ergibt also

$$(19) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(a, |c|) + 12 \operatorname{sign} b \cdot s(d, |b|) \\ = -3 \operatorname{sign} c - 3 \operatorname{sign} b + \dfrac{d}{c} + \dfrac{c}{d} + \dfrac{ad - bc}{dc} + \dfrac{d}{b} + \dfrac{b}{d} + \dfrac{ad - bc}{db}. \end{cases}$$

Ein Vergleich von (17) und (19) zeigt, daß (17) bewiesen sein wird, sobald die Richtigkeit von

$$(20) \qquad \operatorname{sign}\big(c(a+d)\big) + \operatorname{sign}\big(b(a+d)\big) = \operatorname{sign} c + \operatorname{sign} b$$

gezeigt ist. Nun ist (20) klar für $a+d>0$. Für $a+d\leq 0$ und $d>0$ (was wir annehmen durften) folgt $ad<0$, und sömit $bc=ad-1<0$, was entgegengesetzte Vorzeichen von b und c bedeutet. Dann sind beide Seiten von (20) gleich Null. Somit ist (17) bewiesen und damit der Beweis von Satz 7 erbracht.

8. Unter Einführung der Matrix $U=ST$ mit

$$T^2=U^3=1 \tag{21}$$

(wo Faktoren ± 1 vernachlässigt, d.h. die inhomogenen Modulsubstitutionen gemeint sind) kann jede Modulsubstitution *eindeutig* in kürzester Form dargestellt werden[2]) als

$$M=U^{\varepsilon_0}TU^{\varepsilon_1}T\dots TU^{\varepsilon_{\nu+1}}, \tag{22}$$

wo

$$\varepsilon_j=\pm 1, \quad j=1,\dots,\nu$$

$$\varepsilon_0,\varepsilon_{\nu+1}=0 \quad \text{oder} \quad \pm 1.$$

Elemente M, die durch zyklische Vertauschung der Faktoren der Darstellung (22) auseinander hervorgehen, gehören zur selben Klasse. Wenn man das Produkt (22) zyklisch schließt, wird im allgemeinen eine Reduktion durch Anwendung von (21) eintreten. Dann gibt es also in einer Klasse immer eine Transformation von folgender Art

1. die elliptischen T, U, U^{-1}

oder

2. $TU^{\varepsilon_1}TU^{\varepsilon_2}\dots TU^{\varepsilon_\nu}, \quad \varepsilon=\pm 1.$

Hilfssatz 4. *Ist*

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}=TU^{\varepsilon_1}TU^{\varepsilon_2}\dots TU^{\varepsilon_\nu}, \quad \varepsilon_j=\pm 1, \tag{23}$$

dann gilt[3])

$$a<0, \quad d<0, \quad b\geq 0, \quad c\geq 0 \tag{24}$$

Beweis. Diese Behauptungen sind richtig für $\nu=1$, also für

$$TU=TST=\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix},$$

$$TU^{-1}=S^{-1}=\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Angenommen die Behauptungen (24) wären für M in der Form (23) gültig, dann haben wir zu zeigen, daß sie auch für

$$M\cdot TU^{\varepsilon_{\nu+1}}$$

[2]) Den Hinweis auf die Bedeutung der kürzesten Darstellung in diesem Zusammenhang verdanke ich Herrn REIDEMEISTER; vgl. [4], S. 42, 43.

[3]) Möglicherweise zu erreichen durch Multiplikation der Matrix mit -1.

zutreffen. Nun ist

$$M\,T\,U=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}-1 & 0\\ 1 & -1\end{pmatrix}=\begin{pmatrix}-a+b & -b\\ -c+d & -d\end{pmatrix}=\begin{pmatrix}a-b & b\\ c-d & d\end{pmatrix}=\begin{pmatrix}a' & b'\\ c' & d'\end{pmatrix}$$

und

$$M\,T\,U^{-1}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix}=\begin{pmatrix}-a & a-b\\ -c & c-d\end{pmatrix}=\begin{pmatrix}a & b-a\\ c & d-c\end{pmatrix}=\begin{pmatrix}a'' & b''\\ c'' & d''\end{pmatrix}.$$

Aus (24) folgt dann dieselbe Behauptung auch für a', b', c', d', und a'', b'', c'', d'', womit der Hilfssatz bewiesen ist.

9. Nunmehr kann man der Klasseninvariante Ψ eine neue Bedeutung geben. Es gilt nämlich der

Satz 8. *Es sei M in seiner Klasse in der zyklisch reduzierten Form angenommen. Dann gilt:*

$$(25)\qquad \begin{cases}\Psi(T)=\Psi\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}=0,\\ \Psi(U)=\Psi\begin{pmatrix}1 & -1\\ 1 & 0\end{pmatrix}=-2,\quad \Psi(U^{-1})=\Psi\begin{pmatrix}0 & -1\\ 1 & -1\end{pmatrix}=2\end{cases}$$

und im Falle (23)

$$(26)\qquad \Psi(M)=\sum_{j=1}^{\nu}\varepsilon_j.$$

Beweis. Die Formeln (25) liest man direkt aus der Definition (11) von Ψ ab. Die Gl. (26) wird durch die für den Beweis von Hilfssatz 4 angewandte Induktion bewiesen.

Zunächst ist

$$\Psi(T\,U)=\Psi\begin{pmatrix}-1 & 0\\ 1 & -1\end{pmatrix}=1,$$

$$\Psi(T\,U^{-1})=\Psi\begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix}=-1,$$

wie aus (11) folgt. Also stimmt (26) für $\nu=1$.

Es sei nun (26) für ein ν erfüllt, dann bleibt nur zu zeigen, daß

$$(27)\qquad \Psi(M\,T\,U^{\varepsilon_{\nu+1}})=\Psi(M)+\varepsilon_{\nu+1},\qquad \varepsilon_{\nu+1}=\pm 1$$

ist. Sei

$$M=\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

dann ist

$$M\,T\,U=\begin{pmatrix}a-b & b\\ c-d & d\end{pmatrix},$$

also

$$\Psi(M) = \begin{cases} \frac{b}{d} & \text{für} \quad c = 0 \\ \frac{a+d}{c} - 12\,s(d,c) + 3, & c > 0, \end{cases} \tag{28a}$$

$$\Psi(M\,T\,U) = \frac{a-b+d}{c-d} - 12\,s(d, c-d) + 3, \tag{28b}$$

wo nach (24) $c-d=0$ nicht vorkommt, und wo, wieder nach (24), $a+d<0$, $a-b+d<0$ benutzt ist.

Sei zunächst $c=0$, also $a=d=-1$. Dann ist

$$\begin{cases} \Psi(M) = \frac{b}{d} = -b, & \Psi(M\,T\,U) = \frac{-1-b-1}{-d} + 3 \\ & \qquad = -b+1, \end{cases} \tag{29}$$

in Übereinstimmung mit (27).

Sei zweitens $c>0$, also nach (28a), (28b)

$$\Psi(M\,T\,U) - \Psi(M) = \frac{a-b+d}{c-d} - \frac{a+d}{c} - 12\,s(d, c-d) + 12\,s(d,c).$$

Nun ist aber nach (3)

$$-12\,s(d, c-d) = 12\,s(c,d) + 3 - \frac{d^2 + (c-d)^2 + 1}{d\,(c-d)}$$

und somit

$$\Psi(M\,T\,U) - \Psi(M) = 12\,s(c,d) + 12\,s(d,c) + \\ + 3 - \frac{d^2 + (c-d)^2 + 1}{d\,(c-d)} + \frac{a-b+d}{c-d} - \frac{a+d}{c}.$$

Eine nochmalige Anwendung von (3) ergibt dann

$$\Psi(M\,T\,U) - \Psi(M) = \frac{c^2 + d^2 + 1}{c\,d} - \frac{a+d}{c} + \frac{b-c+d}{d} = 1.$$

Diese Gleichung und (29) bestätigen (27) für den Fall $\varepsilon_{\nu+1} = +1$.

Ferner ist

$$\Psi(M\,T\,U^{-1}) = \Psi\begin{pmatrix} a & -a+b \\ c & -c+d \end{pmatrix} = \begin{cases} -\frac{a+b}{d} & c=0 \\ \frac{a-c+d}{c} - 12\,s(d,c) + 3, & c>0 \end{cases} \tag{30}$$

und somit für $c=0$, $a=d=-1$:

$$\Psi(M) = -b, \quad \Psi(M\,T\,U^{-1}) = -1-b,$$

während sich für $c>0$ aus (28a) und (30)

$$\Psi(M\,T\,U^{-1}) - \Psi(M) = -\frac{c}{c} = -1$$

ergibt. Damit ist (27) auch für den Fall $\varepsilon_{\nu+1}=-1$ bewiesen und der Induktionsbeweis von Satz 8 zu Ende geführt.

10. Im allgemeinen kann man nur aussagen, daß $\Psi(M_1M_2)$ und $\Psi(M_1)+\Psi(M_2)$ sich nur um Vielfache von 3 unterscheiden können. Jedoch gilt der

Satz 9. *Wenn M keine elliptische Substitution ist, dann ist für ganzes k*

$$\Psi(M^k)=k\,\Psi(M).$$

Beweis. Dies folgt zunächst für $k>0$ daraus, daß in dem zyklischen Zusammenschluß von M_k k mal die gleichen Reduktionen auftreten, die nur einmal bei dem zyklischen Zusammenschluß von M auftreten. Also besteht das zyklisch reduzierte M_k aus k Wiederholungen des zyklisch reduzierten M, was den Satz für $k>0$ beweist. Für $k<0$ folgt der Satz dann aus der zweiten Gl. (12), und für $k=0$ aus $\Psi(E)=0$, wo E die Einheitsmatrix ist.

In etwas anderer Richtung liegt der

Satz 10. *Wenn für*

$$M_1=\begin{pmatrix}a_1 & b_1\\ c_1 & d_1\end{pmatrix},\qquad M_2=\begin{pmatrix}a_2 & b_2\\ c_2 & d_2\end{pmatrix}, \tag{31}$$

die Ungleichungen

$$\begin{cases} a_1<0,\ d_1<0, & c_1>0,\ b_1>0\\ a_2<0,\ d_2<0, & c_2>0,\ d_2>0\end{cases} \tag{32}$$

gelten, so ist

$$\Psi(M_1M_2)=\Psi(M_1)+\Psi(M_2). \tag{33}$$

Beweis. Die Ungleichungen (32) sorgen dafür, daß in der Normaldarstellung (22) M_1 so endet und M_2 so anfängt, daß keine gegenseitigen Reduktionen in dem Produkt M_1M_2 stattfinden.

Um dies einzusehen, ändern wir die Bedingungen zu (22) dahin ab, daß wir nur nicht-negative Potenzen zulassen, was dadurch geschieht, daß jeder etwa auftretende Exponent -1 durch 2 ersetzt wird. Dann werde die Definition $U=ST$ benutzt, so daß (22) umgeformt wird in

$$\begin{cases} M=(ST)^{\eta_0}T(ST)^{\eta_1}T\dots(ST)^{\eta_\nu}T(ST)^{\eta_{\nu+1}},\\ \eta_0,\eta_{\nu+1}=0,1,2,\quad \eta_j=1,2,\quad j=1,\dots,\nu.\end{cases} \tag{34}$$

Dann und nur dann, wenn $\eta_0=0$, $\eta_{\nu+1}>0$ wird M in (22) die spezielle Form (23) haben, d.h. mit T anfangen und mit einer Potenz von U enden. Wir wollen zeigen, daß aus

$$a<0,\quad d<0,\quad c>0 \tag{35}$$

folgt $\eta_0=0$, $\eta_{\nu+1}>0$.

In der Tat, (34) kann nach gehörigem Zusammenziehen auf Grund von $T^2=1$ geschrieben werden:

$$M=S^{q_0}TS^{q_1}\dots TS^{q_l}TS^{q_{l+1}},\quad q_j\geq 1,\ j=1,\dots l,\quad q_0\geq 0,\ q_{l+1}\geq 0, \tag{36}$$

wo $q_0 > 0$ aus $\eta_0 > 0$, und $q_{l+1} > 0$ aus $\eta_{\nu+1} = 0$ folgen würde. Man sieht ferner leicht, daß wegen (34) von den nicht-verschwindenden q_j nur das erste und letzte gleich 1 sein können, während für die übrigen sogar $q_j \geq 2$ gilt. Nach (36) kann man

$$M(\tau) = \frac{a\tau + b}{c\tau + d}$$

als einen Kettenbruch

$$\frac{a\tau + b}{c\tau + d} = q_0 - \cfrac{1}{q_1 - \ddots - \cfrac{1}{q_l - \cfrac{1}{q_{l+1} + \tau}}}$$

schreiben, woraus durch $\frac{1}{\tau} \to 0$

$$\frac{a}{c} = q_0 - \cfrac{1}{q_1 - \ddots - \cfrac{1}{q}}$$

folgt, so daß $q_0 > 0$

$$\frac{a}{c} \geq 0$$

nach sich ziehen würde, gegen die Ungleichungen (35). Ebenso beweist man aus

$$M^{-1} = S^{-q_{l+1}} T S^{-q} \dots T S^{-q_1} T,$$

also aus

$$-\frac{d}{c} = -q_{l+1} - \cfrac{1}{-q_l - \ddots - \cfrac{1}{-q_1}},$$

daß $q_{l+1} > 0$ die Ungleichung

$$-\frac{d}{c} \leq 0$$

zur Folge hätte, wieder gegen (35). Es ist also $q_0 = q_{l+1} = 0$ und folglich $\eta_0 = 0, \eta_{\nu+1} > 0$, und M hat die Form (23), wobei nur $\varepsilon_{\nu+1}$ statt ε_ν zu schreiben ist.

Daher haben infolge von (32) die Substitutionen M_1 und M_2 auch die Form (23). Dann aber tritt in dem Produkt $M_1 M_2$ keine Reduktion ein, da auf das letzte $U^{\prime\nu}$ von M, das erste T von M_2 folgt. Damit ist Satz 10 bewiesen.

Wenn man in (33) die Matrizen von (31) einträgt und die Definition (11) von Ψ aus Satz 7 heranzieht, so folgt nach einigen Rechnungen das

Korollar. *Unter den Bedingungen* (32) *ist*

$$s(d, |c|) + s(d_1, c_1) + s(d_2, c_2) = \frac{1}{4} - \frac{c_1^2 + c_2^2 + c^2}{12 c_1 c_2 |c|}$$

mit

$$c = c_1 a_2 + d_1 c_2, \quad d = c_1 b_2 + d_1 d_2.$$

Diese Formel könnte auch analytisch aus der Theorie der Funktion $\log \eta(\tau)$ gewonnen werden. Sie folgt übrigens auch aus der dreigliedrigen Formel (30) in [2].

III

11. Für die elliptischen Substitutionen, die mit T oder U oder U^{-1} äquivalent sind, also nur zu 3 Klassen gehören können, ist die Spur $a+d$ gleich 0 oder ± 1. Wenn man die Spur $a+d=m$ mit $|m| \geq 2$ vorgibt, wobei man $a+d$ negativ wählen kann, so kann man nach dem Hilfssatz 4 und den ihm vorausgehenden Erörterungen in derselben Klasse stets eine Substitution in der Form (23) finden, für die $a<0$, $d<0$ ist. Solche a, d gibt es aber bei festem $a+d=m$ nur endlich viele.

Ist $m \leq -3$, so ist $a \cdot d \neq 1$, also

$$bc = ad - 1 \neq 0,$$

und von den b, c (die beide nach Hilfssatz 4 positiv zu sein haben), gibt es daher wieder nur endlich viele. Also gibt es mit vorgegebener Spur $a+d=m \leq -3$ nur je endlich viele Klassen von Modulsubstitutionen.

Für $a+d=-2$, also $a=d=-1$, $bc=0$ gibt es die unendlich vielen Substitutionen

$$(TU)^c = \begin{pmatrix} -1 & 0 \\ c & -1 \end{pmatrix}, \quad c>0,$$

$$(TU^{-1})^b = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix}, \quad b>0,$$

die verschiedene Normalformen (23) zeigen, und also zu verschiedenen Klassen gehören. Dies sind die parabolischen Substitutionen. Es ist übrigens

$$\Psi\left((TU)^c\right) = c, \quad \Psi\left((TU^{-1})^b\right) = -b.$$

12. Einen anderen Einblick in die Verteilung auf die Klassen bei gegebenem $m=a+d$ gewinnt man durch eine Zuordnung von quadratischen Formen zu den Modulsubstitutionen.

Wir gehen dazu wieder zu den homogenen Modulsubstitutionen über, unterscheiden also die unimodularen Matrizen M und $-M$. Dann werde der Modulsubstitution

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

die quadratische Form

$$Q(x, y) = c x^2 + (d-a) x y - b y^2 \tag{36}$$

mit der Diskriminante

$$\Delta = (d-a)^2 + 4bc = (a+d)^2 - 4 = m^2 - 4$$

zugeordnet. Es ist klar, daß zu

$$-M^{-1}=\begin{pmatrix}-d & b\\ c & -a\end{pmatrix}$$

dieselbe Form Q gehört wie zu M.

Umgekehrt, hat

(37) $$Q(x,y)=Ax^2+Bxy+Cy^2$$

die Diskriminante

(38) $$\Delta=B^2-4AC=m^2-4,$$

dann gehört sie zu

$$M=\begin{pmatrix}\frac{1}{2}(m-B) & -C\\ A & \frac{1}{2}(m+B)\end{pmatrix} \text{ und } -M^{-1}=\begin{pmatrix}-\frac{1}{2}(m+B) & -C\\ A & \frac{1}{2}(m-B)\end{pmatrix}.$$

Die Elemente in M sind ganz, da aus (38)

$$B\equiv m \pmod 2$$

folgt.

Im allgemeinen liegen M und $-M^{-1}$ in verschiedenen Ähnlichkeitsklassen der homogenen Modulgruppe, da sie verschiedene Spuren $a+d$ und $-a-d$ haben. In derselben Klasse können sie nur für $a+d=0$ liegen, und tun es auch wirklich, weil dann $M=-M^{-1}$.

Der Ähnlichkeitsklasseneinteilung der Modulgruppe entspricht nun die Einteilung der quadratischen Formen in eigentliche Äqúivalenzklassen. Wenn die Koeffizienten der quadratischen Form (37) die Matrix

$$Q=\begin{pmatrix}A & \frac{B}{2}\\ \frac{B}{2} & C\end{pmatrix}$$

bilden, so handelt es sich also um den folgenden

Satz 11. *Wenn der quadratischen Form (oder ihrer Koeffizientenmatrix) Q die Modulsubstitution M (zusammen mit $-M^{-1}$) zugeordnet ist, dann gehört zu $L'QL$ die Matrix $L^{-1}ML$ (zusammen mit $-L^{-1}M^{-1}L$), wo L eine beliebige Modulsubstitution ist und L' die Transponierte von L bedeutet.*

Beweis. Der Satz braucht nur für $L=S$ und $L=T$ bewiesen zu werden.

Es sei

$$M=\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

dann ist [s. (13a)]

$$S^{-1}MS=\begin{pmatrix}a' & b'\\ c' & d'\end{pmatrix}=\begin{pmatrix}a-c & a-c+b-d\\ c & c+d\end{pmatrix}.$$

Zu M gehört nach (36) die Koeffizientenmatrix

$$Q=\begin{pmatrix}c & \frac{d-a}{2}\\ \frac{d-a}{2} & -b\end{pmatrix},$$

und in der Tat gilt dann

$$S'\,Q\,S=\begin{pmatrix} c & c+\frac{d-a}{2} \\ c+\frac{d-a}{2} & c-b+d-a \end{pmatrix}=\begin{pmatrix} c' & \frac{d'-a'}{2} \\ \frac{d'-a'}{2} & -b' \end{pmatrix},$$

wie es sein muß. Ebenso ist [s. (14a)]

$$T^{-1}M\,T=\begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}=\begin{pmatrix} d & -e \\ -b & a \end{pmatrix}$$

und

$$T'\,Q\,T=\begin{pmatrix} -b & \frac{a-d}{2} \\ \frac{a-d}{2} & c \end{pmatrix}=\begin{pmatrix} c'' & \frac{d''-a''}{2} \\ \frac{d''-a''}{2} & -b'' \end{pmatrix},$$

wie es der zu beweisende Satz verlangt.

(Bemerkung. LATIMER und MACDUFFEE [5] und O. TAUSSKY [6] haben eine Zuordnung von Matrizen n-ter Ordnung zu Idealen in algebraischen Körpern n-ten Grades studiert. Wir ziehen hier statt der Ideale in quadratischen Körpern die quadratischen Formen vor, die auch für $\Delta=0$ brauchbar sind.)

13. Satz 7 zusammen mit Satz 11 ergeben nun den

Satz 12. *Für die quadratischen Formen $Q(x,y)=Ax^2+Bxy+Cy^2$ mit der Diskriminante $\Delta=B^2-4AC=m^2-4$ bildet die Funktion*

$$\Psi\begin{pmatrix} \frac{1}{2}(m-B) & -c \\ A & \frac{1}{2}(m+B) \end{pmatrix}$$

eine Klasseninvariante.

Das folgende Beispiel ist lehrreich: Für $m=45$, $\Delta=45^2-4=43\cdot 47=2021$ gehören die beiden Formen

$$x^2+xy-505y^2 \quad \text{und} \quad 17x^2-7xy-29y^2$$

verschiedenen Klassen an, denn für die erstere findet man

$$\Psi\begin{pmatrix} 22 & 505 \\ 1 & 23 \end{pmatrix}=42$$

und für die letztere

$$\Psi\begin{pmatrix} 19 & 29 \\ 17 & 26 \end{pmatrix}=-6.$$

Die beiden Formen gehören aber demselben Geschlecht an, weil

$$\left(\frac{A}{43}\right)=+1, \quad \left(\frac{A}{47}\right)=+1$$

für $A=1$ und $A=17$ ist[4]). Die Funktion Ψ ist also keine generische Invariante.

[4]) Ein anderes Beispiel dieser Art, für $\Delta=25^2-4=621$ wurde mir von Herrn R. Ayoub mitgeteilt.

14. Man kann nun umgekehrt den Satz 12 benutzen, um in vielen Fällen $s(h, k)$ ohne eine auf Gl. (3) gegründete Rekursion zu finden, nämlich zunächst, wenn die Klassenzahl zu $\Delta = m^2 - 4$ gleich 1 ist. Dies trifft nur für $|m| = 0, 1, 3$, also für $\Delta = -4, -3, 5$ zu. Da für diese m alle unimodularen Substitutionen der Spur $a + d = m$ zur gleichen Klasse gehören müssen, so haben wir hier

$$\Psi\begin{pmatrix} a & b \\ c & m-a \end{pmatrix} = \Psi\begin{pmatrix} m & -1 \\ 1 & 0 \end{pmatrix}$$

mit $a(m-a) - bc = 1$, $m = 0, \pm 1, \pm 3$, wobei wir $c > 0$ nehmen können. Wenn man in diese Gleichung die Werte aus (11) einsetzt, so erhält man den

Satz 13. *Es ist für* $c > 0$ *und* $a(m-a) \equiv 1 \pmod{c}$

$$s(a, c) = \frac{(1-c)\,m}{12\,c}$$

für $m = 0, \pm 1, \pm 3$.

Dieser Satz besagt für $m = 0$, also $a^2 \equiv -1 \pmod{c}$ nur einen Teil von Satz 1. Für ungerades m aber ist $a(m-a)$ gerade, also c ungerade. Dann lassen sich die Aussagen für $m = \pm 1, \pm 3$ auch folgendermaßen schreiben:

Für ungerades $c > 0$ und $(2a \pm 1)^2 \equiv -3 \pmod{c}$ ist

$$s(a, c) = \pm \frac{c-1}{12c}; \tag{39}$$

und für $(2a \pm 3)^2 \equiv 5 \pmod{c}$ ist

$$s(a, c) = \pm \frac{c-1}{4c}. \tag{40}$$

Dieser Fall ist besonders bemerkenswert, da er zu $\Delta > 0$ gehört, wo $M(\tau)$ hyperbolisch ist und nur Fixpunkte auf der reellen Achse hat. Ein Beispiel für (40) ist

$$a = 16,\ c = 61,\ m = +3,\ (32+3)^2 = 1225 \equiv 5 \pmod{61},$$

also

$$s(16, 61) = \frac{60}{4 \cdot 61} = \frac{15}{61}.$$

Auch wenn die Klassenzahl von Δ nicht 1 aber klein ist, könnte man Ergebnisse erzielen. Zum Beispiel für $m = 4$, $\Delta = 12$ gibt es zwei Klassen quadratischer Formen, von denen

$$x^2 - 4xy + y^2 \quad \text{und} \quad -x^2 + 4xy - y^2$$

Repräsentanten sind. Diesen Formen sind beziehungsweise die unimodularen Matrizen

$$\begin{pmatrix} 4 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 1 \\ -1 & 4 \end{pmatrix}$$

zugeordnet, für die man erhält:

$$\Psi\begin{pmatrix} 4 & -1 \\ 1 & 0 \end{pmatrix} = 1, \quad \Psi\begin{pmatrix} 0 & 1 \\ -1 & 4 \end{pmatrix} = -1.$$

Wir haben also mit $c>0$, $a(4-a)-bc=1$

$$\Psi\begin{pmatrix} a & b \\ c & 4-a \end{pmatrix} = \pm 1$$

oder

$$s(a,c) = \frac{4-3c\mp c}{12c}. \tag{41}$$

Sei zunächst $3 \nmid c$. Dann wissen wir nach Hilfssatz 2, daß 3 kein Teiler des Nenners von $s(a,c)$ sein kann, so daß also folgt

$$s(a,c) = \begin{cases} \frac{4-4c}{12c} = \frac{1-c}{3c} & c \equiv 1 \pmod 3, \\ \frac{4-2c}{12c} = \frac{2-c}{6c} & c \equiv 2 \pmod 3. \end{cases} \tag{42}$$

Ist ferner $3/c$, so zeigt die Bedingung $a(4-a) \equiv 1 \pmod c$ oder

$$(a-2)^2 \equiv 3 \pmod c, \tag{43}$$

daß $9 \nmid c$, also nur $c \equiv \pm 3 \pmod 9$ in Frage kommen.

Nun folgt aus (41)

$$12c\, s(a,c) \equiv 4 \mp c \pmod{3c},$$

anderseits aus (7), Hilfssatz 3,

$$12ac\, s(a,c) \equiv a^2+1 \pmod{3c},$$

also

$$4a \mp ac \equiv a^2+1 \pmod 9.$$

Wegen (43) ist

$$a-2 \equiv 0 \pmod 3, \quad (a-2)^2 \equiv 0 \pmod 9,$$

also

$$3 \mp ac \equiv 0 \pmod 9,$$

oder

$$3 \pm c \equiv 0 \pmod 9,$$

so daß, da die oberen und unteren Vorzeichen hier und in (41) sich entsprechen,

$$s(a,c) = \begin{cases} \frac{4-4c}{12c} & c \equiv -3 \pmod 9 \\ \frac{4-2c}{12c} & c \equiv 3 \pmod 9. \end{cases}$$

Zusammenfassend haben wir also

Satz 14. *Wenn* $c>0$, $(a-2)^2 \equiv 3 \pmod c$, *so ist*

$$s(a,c) = \begin{cases} \frac{1-c}{3c} & \textit{für } c \equiv 1 \pmod 3 \textit{ und für } c \equiv -3 \pmod 9, \\ \frac{2-c}{6c} & \textit{für } c \equiv -1 \pmod 3 \textit{ und für } c \equiv 3 \pmod 9. \end{cases}$$

Von dieser Art könnte man noch weitere Sätze für geeignetes $\Delta > 0$ aufstellen.

15. Es bleibt noch der parabolische Fall $\Delta = 0$, $m = \pm 2$ zu diskutieren. Hier haben wir, wie schon in § 11 erwähnt ist, unendlich viele Klassen.

Sei zunächst

$$a + d = 2, \tag{44}$$

also

$$(d-1)^2 = -bc.$$

Zu

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

gehört die quadratische Form

$$Q(x,y) = cx^2 + 2(d+1)xy - by^2$$
$$= \frac{1}{c}(cx + (d-1)y)^2.$$

Wenn

$$\lambda = (c, d-1), \tag{45}$$
$$c = \lambda\gamma, \quad d - 1 = \lambda\delta$$

gesetzt wird, erhalten wir

$$Q(x,y) = \frac{\lambda^2}{c}(\gamma x + \delta y)^2.$$

Hier ist einerseits λ^2/c ganz, denn

$$\lambda^2 = (c^2, (d-1)^2)$$

zugleich mit

$$(d-1)^2 \equiv 0 \pmod{c}. \tag{46}$$

Anderseits ist

$$(\gamma, \delta) = 1,$$

also lassen sich α, β finden, so daß

$$x' = \alpha x + \beta y$$
$$y' = \gamma x + \delta y$$

eine Modulsubstitution darstellt, und somit ist

$$Q(x,y) \sim \frac{\lambda^2}{c} y^2.$$

Daher gilt innerhalb der Modulgruppe die Äquivalenz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} 1 & -\frac{\lambda^2}{c} \\ 0 & 1 \end{pmatrix},$$

so daß also

$$\Psi\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \Psi\begin{pmatrix} 1 & -\frac{\lambda^2}{c} \\ 0 & 1 \end{pmatrix} = -\frac{\lambda^2}{c}$$

ist. Ausführlich heißt dies, unter der Annahme $c>0$ (sonst könnte man die Vorzeichen umkehren und $a+d=-2$ betrachten),

$$\frac{2}{c}-12s(a,c)-3=-\frac{\lambda^2}{c}, \tag{47}$$

unter der Bedingung (46).

Wenn man endlich statt (44) die Bedingung

$$a+d=-2$$

nimmt, so werden (45), (46) durch

$$\lambda=(c,d+1),$$
$$(d+1)^2\equiv 0 \pmod{c}$$

ersetzt, und statt (47) erscheint, $c>0$ vorausgesetzt,

$$-\frac{2}{c}-12s(a,c)+3=\frac{\lambda^2}{c}.$$

Schließlich beachten wir noch $s(a,c)=s(d,c)$ und erhalten somit den

Satz 15. *Ist* $c>0$, $(d\pm 1)^2\equiv 0 \pmod{c}$, *so ist*

$$s(d,c)=\mp\frac{\lambda^2+2-3c}{12c},$$

mit $\lambda=(c,d\pm 1)$, *wobei die unteren und die oberen Vorzeichen zusammengehören.*

Für quadratfreies c ist stets $\lambda=c$, und der Satz besagt dann nicht mehr als (15). Er enthält jedoch eine neue Aussage für c mit quadratischem Faktor. Wir erhalten z.B. für

$$c=50,\quad d=10j\pm 1,\quad j=1,2,3,4,\quad \text{also}\quad \lambda=10$$
$$s(10j\pm 1,50)=\pm\frac{100+2-150}{600}=\mp\frac{2}{25}.$$

Literatur

[1] Mordell, L. J.: Lattice points in a tetrahedron and generalized Dedekind sums. J. Indian Math. Soc. **15**, 41—46 (1951). — [2] Rademacher, H.: Generalization of the reciprocity formula for Dedekind sums. Duke Math. J. **21**, 391—397 (1954). — [3] Rademacher, H., and A. Whiteman: Theorems on Dedekind sums. Amer. Math. J. **63**, 377—407 (1941). — [4] Reidemeister, K.: Einführung in die kombinatorische Topologie. Braunschweig 1932. — [5] Latimer, C. G., and C. C. MacDuffee: A correspondence between classes of ideals and classes of matrices. Ann. of Math. (2) **34**, 313—316 (1933). — [6] Taussky, O.: On a theorem of Latimer and MacDuffee. Canad. J. Math. **1**, 300—302 (1949).

Prof. H. Rademacher, University of Pennsylvania, Philadelphia 4, Pa. (USA)

(Eingegangen am 24. August 1955)

STÖHR, A.
Math. Zeitschr. Bd. 63, S. 464—477 (1956)

Zur Funktionentheorie im Raum der symmetrischen Matrizen

Dem Andenken meines verehrten Lehrers I. SCHUR gewidmet

Von

ALFRED STÖHR

Inhalt

Seite
§ 1. Fragestellungen 464
§ 2. Rangerniedrigende und rangerhaltende lineare Funktionen 467
§ 3. Differentiell rangerniedrigende und differentiell spezielle Funktionen 471
§ 4. Die Eigenwerte bei rangerniedrigender linearer Abbildung 473
§ 5. Ein Analogon zum SCHWARZschen Lemma 475

§ 1. Fragestellungen

In der Funktionentheorie einer Veränderlichen ist die Gesamtheit der Funktionen, die den Einheitskreis $z\bar{z}<1$ konform auf sich abbilden, bekanntlich durch

$$\eta\frac{z-\alpha}{-\bar{\alpha}z+1} \qquad (|\alpha|<1,\ |\eta|=1) \tag{1}$$

gegeben. Bei jeder solchen Abbildung bleiben nichteuklidische Abstände invariant. Für jede im Einheitskreis reguläre Funktion $f(z)$, die dort der Bedingung $|f(z)|<1$ genügt, die also den Einheitskreis in sich abbildet, gilt dann für je zwei Punkte z, z_1 aus dem Einheitskreis die Ungleichung

$$\delta(f(z), f(z_1)) \leq \delta(z, z_1), \tag{2}$$

wo δ den nichteuklidischen Abstand bezeichnet. Für den Nachweis dieser Tatsache kann man sich, indem man Hilfsabbildungen von der Gestalt (1) heranzieht, ohne Beschränkung der Allgemeinheit auf den Spezialfall $z_1=f(z_1)=0$ beschränken. Da die nichteuklidische Entfernung $\delta(z, 0)$ eine monotone Funktion von $|z|$ ist, läuft in diesem Spezialfall der Nachweis von (2) auf das SCHWARZsche Lemma

$$|f(z)| \leq |z| \tag{3}$$

hinaus. Aus (3) folgt

$$|f'(0)| \leq 1, \tag{4}$$

woraus umgekehrt nach Anwendung von Transformationen (1) die Ungleichung

$$\frac{|df(z)|}{1-|f(z)|^2} \leq \frac{|dz|}{1-|z|^2}$$

und daraus durch Integration wieder (2) und (3) hergeleitet werden können.

Ziel der folgenden Untersuchungen ist es, die genannten Sätze auf gewisse Funktionen auszudehnen, deren unabhängige und abhängige Veränderliche n-reihige bzw. m-reihige symmetrische Matrizen sind; die Elemente der abhängigen Matrix sollen dabei analytisch von den Elementen der unabhängigen Matrix abhängen. Beispiele solcher Funktionen erhält man für $m=n$ wie folgt: $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ seien n-reihige quadratische Matrizen. $\begin{pmatrix}\mathfrak{A} & \mathfrak{B}\\ \mathfrak{C} & \mathfrak{D}\end{pmatrix}$ heiße symplektisch, wenn $\mathfrak{A}'\mathfrak{C}=\mathfrak{C}'\mathfrak{A}$, $\mathfrak{B}'\mathfrak{D}=\mathfrak{D}'\mathfrak{B}$, $\mathfrak{A}'\mathfrak{D}-\mathfrak{C}'\mathfrak{B}=\mathfrak{E}$ ist. Dann führt die Abbildung

$$(5)\qquad \mathfrak{Z}\to\mathfrak{W}=(\mathfrak{A}\mathfrak{Z}+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})^{-1}\quad(\text{für } \det(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})\neq 0)$$

symmetrische Matrizen $\mathfrak{Z}$ in symmetrische Matrizen $\mathfrak{W}$ über; wir nennen $\mathfrak{W}$ eine symplektische linear-gebrochene Funktion von $\mathfrak{Z}$.

Sind $\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3, \mathfrak{Z}_4$ irgend vier symmetrische Matrizen, so bilde man mit ihnen den Ausdruck

$$\mathfrak{R}=\mathfrak{R}(\mathfrak{Z}_1,\mathfrak{Z}_2,\mathfrak{Z}_3,\mathfrak{Z}_4)=(\mathfrak{Z}_1-\mathfrak{Z}_3)(\mathfrak{Z}_1-\mathfrak{Z}_4)^{-1}(\mathfrak{Z}_2-\mathfrak{Z}_4)(\mathfrak{Z}_2-\mathfrak{Z}_3)^{-1}$$

(falls die rechte Seite einen Sinn hat). Es empfiehlt sich, nicht nur diese Matrix, sondern auch jede der gleichen Ähnlichkeitsklasse angehörige Matrix $\mathfrak{M}\mathfrak{R}\mathfrak{M}^{-1}$ als Doppelverhältnis von $\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3, \mathfrak{Z}_4$ zu bezeichnen; die (bekanntlich nur von der Ähnlichkeitsklasse abhängenden) Eigenwerte von $\mathfrak{R}$ nennen wir die Eigenwerte des Doppelverhältnisses[1]). Die durch $\mathfrak{R}$ repräsentierte Ähnlichkeitsklasse ist gegenüber symplektischen linear-gebrochenen Substitutionen invariant, was man durch mehrmalige Anwendung der Formel

$$\begin{aligned}\mathfrak{W}_i-\mathfrak{W}_k&=(\mathfrak{A}\mathfrak{Z}_i+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}_i+\mathfrak{D})^{-1}-(\mathfrak{A}\mathfrak{Z}_k+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}_k+\mathfrak{D})^{-1}\\&=(\mathfrak{A}\mathfrak{Z}_i+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}_i+\mathfrak{D})^{-1}-(\mathfrak{Z}_k\mathfrak{C}'+\mathfrak{D}')^{-1}(\mathfrak{Z}_k\mathfrak{A}'+\mathfrak{B}')\\&=(\mathfrak{C}\mathfrak{Z}_k+\mathfrak{D})'^{-1}(\mathfrak{Z}_i-\mathfrak{Z}_k)(\mathfrak{C}\mathfrak{Z}_i+\mathfrak{D})^{-1}\end{aligned}$$

erkennt.

Die folgenden Definitionen und Sätze entnehmen wir einer Arbeit von SIEGEL [2]. — Im Raum der symmetrischen Matrizen $\mathfrak{Z}$ werde der Einheitskreis durch

$$(6)\qquad \mathfrak{E}-\mathfrak{Z}\overline{\mathfrak{Z}}>0$$

definiert. Für zwei im Einheitskreis gelegene Matrizen $\mathfrak{Z}, \mathfrak{Z}_1$ bilde man

$$\mathfrak{R}=\mathfrak{R}(\mathfrak{Z},\overline{\mathfrak{Z}}^{-1},\mathfrak{Z}_1,\overline{\mathfrak{Z}}_1^{-1})=(\mathfrak{Z}-\mathfrak{Z}_1)(\mathfrak{Z}-\overline{\mathfrak{Z}}_1^{-1})^{-1}(\overline{\mathfrak{Z}}^{-1}-\overline{\mathfrak{Z}}_1^{-1})(\overline{\mathfrak{Z}}^{-1}-\mathfrak{Z}_1)^{-1}.$$

Durch die Umformung

$$(7)\qquad \mathfrak{R}=(\mathfrak{Z}-\mathfrak{Z}_1)(\mathfrak{E}-\overline{\mathfrak{Z}}_1\mathfrak{Z})^{-1}(\overline{\mathfrak{Z}}-\overline{\mathfrak{Z}}_1)(\mathfrak{E}-\mathfrak{Z}_1\overline{\mathfrak{Z}})^{-1}$$

wird dieser Ausdruck auch für Matrizen mit verschwindender Determinante sinnvoll. [Für $\mathfrak{Z}_1=0$ nimmt (7) die einfache Gestalt $\mathfrak{R}=\mathfrak{Z}\overline{\mathfrak{Z}}$ an.] Zu (7)

[1]) Über Doppelverhältnisse vgl. auch [1] sowie die dort angegebene Literatur.

bestimme man eine Matrix $\mathfrak{T}$ mittels der konvergenten Reihe

$$\mathfrak{T} = 4\mathfrak{R}\left(\sum_{k=0}^{\infty} \frac{\mathfrak{R}^k}{2k+1}\right)^2$$

(also $\mathfrak{T}\mathfrak{g}\frac{\sqrt{\mathfrak{T}}}{2} = \sqrt{\mathfrak{R}}$) und setze dann $s = \sqrt{\text{Spur}\,\mathfrak{T}}$. Die dadurch definierte nichtnegative reelle Zahl s wollen wir hier als die symplektische Entfernung der Matrizen $\mathfrak{Z}, \mathfrak{Z}_1$ im Einheitskreis bezeichnen. Ist ein im Einheitskreis (6) gelegenes rektifizierbares Kurvenstück gegeben, so nennen wir das darüber erstreckte Integral

$$\int ds = 2\int \sqrt{\text{Spur}\{d\mathfrak{Z}(\mathfrak{E} - \overline{\mathfrak{Z}}\mathfrak{Z})^{-1} d\overline{\mathfrak{Z}}(\mathfrak{E} - \mathfrak{Z}\overline{\mathfrak{Z}})^{-1}\}} \tag{8}$$

die symplektische Bogenlänge im Einheitskreis. Jede analytische Abbildung des Einheitskreises (6) auf sich wird durch eine symplektische linear-gebrochene Funktion bewirkt. Bei jeder solchen Abbildung bleibt die durch (7) repräsentierte Ähnlichkeitsklasse und folglich auch die symplektische Entfernung s und die symplektische Bogenlänge (8) invariant. Zu je zwei Punkten $\mathfrak{Z}_1, \mathfrak{W}_1$ im Einheitskreis gibt es eine derartige Abbildung, die $\mathfrak{Z}_1$ in $\mathfrak{W}_1$ überführt. Die Gesamtheit aller Abbildungen, die den Einheitskreis festlassen und die Nullmatrix in die Nullmatrix überführen, wird durch

$$\mathfrak{W} = \mathfrak{Z}[\mathfrak{U}] \qquad (\mathfrak{U} \text{ unitär})$$

gegeben[2]). Jede dem Einheitskreis angehörende symmetrische Matrix $\mathfrak{M}$ kann in der Gestalt

$$\mathfrak{M} = \mathfrak{P}[\mathfrak{U}] \qquad (\mathfrak{U} \text{ unitär}, \mathfrak{P} \text{ Diagonalmatrix}) \tag{9}$$

geschrieben werden, wobei die Diagonalelemente $p_1, \ldots, p_n$ von $\mathfrak{P}$ beliebig gewählt werden können bis auf die Einschränkung, daß $|p_1|^2, \ldots, |p_n|^2$ die Eigenwerte der hermiteschen Matrix $\mathfrak{M}\overline{\mathfrak{M}}$ sein müssen. — Soweit die Anknüpfung an die erwähnte Arbeit von SIEGEL.

Unter den analytischen Funktionen $\mathfrak{W} = f(\mathfrak{Z})$ werden wir in § 3 diejenigen als „differentiell rangerniedrigend" aussondern, bei denen stets $\text{Rang}\, d\mathfrak{W} \leq \text{Rang}\, d\mathfrak{Z}$ ist; zu dieser Klasse von Funktionen gehören insbesondere die symplektischen linear-gebrochenen Funktionen. Eines unserer Ergebnisse wird dann lauten, daß bei Abbildungen $\mathfrak{Z} \to \mathfrak{W} = f(\mathfrak{Z})$, die durch eine differentiell rangerniedrigende Funktion bewirkt werden und die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes hinein abbilden, alle symplektischen Entfernungen nicht zunehmen. Der Beweis erfolgt so, daß der Satz zunächst für unendlich kleine Entfernungen bewiesen wird; für diesen Fall wird in § 4 noch eine schärfere, die Eigenwerte des Doppelverhältnisses betreffende Aussage hergeleitet.

Solche Funktionen im Raum der symmetrischen Matrizen sind für gewisse elektrotechnische Probleme von Interesse. In der Theorie der linearen Schaltungen wird der Widerstand eines passiven $2n$-Pols durch eine n-reihige

[2]) Wir benutzen die Abkürzung $\mathfrak{A}[\mathfrak{B}] = \mathfrak{B}'\mathfrak{A}\mathfrak{B}$.

quadratische Matrix dargestellt. Durch eine passive lineare Schaltung mit $2m$ Eingangsklemmen und $2n$ Ausgangsklemmen wird der $2n$-Pol in einen $2m$-Pol transformiert. In einigen wichtigen Fällen bedeutet das für die Widerstände eine Abbildung von symmetrischen n-reihigen Matrizen mit positivem Realteil auf symmetrische m-reihige Matrizen mit positivem Realteil, die überdies durch eine differentiell rangerniedrigende Funktion bewirkt wird. Die symmetrischen Matrizen $\mathfrak{M}$ mit positivem Realteil lassen sich aber durch die symplektische linear-gebrochene Funktion

$$\mathfrak{Z} = (\mathfrak{M} - \mathfrak{E})\,(\mathfrak{M} + \mathfrak{E})^{-1}$$

umkehrbar eindeutig auf die Matrizen $\mathfrak{Z}$ im Einheitskreis abbilden. Bei dieser Abbildung bleiben die [in der oberen Halbebene mittels

$$\mathfrak{R} = (\mathfrak{M} - \mathfrak{M}_1)\,(\mathfrak{M} + \overline{\mathfrak{M}}_1)^{-1}(\overline{\mathfrak{M}} - \overline{\mathfrak{M}}_1)\,(\overline{\mathfrak{M}} + \mathfrak{M}_1)^{-1}$$

zu definierenden] symplektischen Entfernungen erhalten, so daß die dafür im Einheitskreis gewonnenen Sätze auch auf die Matrizen mit positivem Realteil Anwendung finden können.

§ 2. Rangerniedrigende und rangerhaltende lineare Funktionen

Die Buchstaben $\mathfrak{W}$ und $\mathfrak{Z}$ bezeichnen im folgenden stets symmetrische Matrizen; $\mathfrak{W}$ sei dabei m-reihig, $\mathfrak{Z}$ sei n-reihig.

Die Begriffe „rangerniedrigend" und „rangerhaltend" mögen zunächst für lineare homogene Funktionen erklärt werden. Die Elemente der Matrix $\mathfrak{W}$ mögen linear und homogen von den Elementen der Matrix $\mathfrak{Z}$ abhängen. Die Funktion $\mathfrak{W} = f(\mathfrak{Z})$ und die dadurch bewirkte Abbildung $\mathfrak{Z} \to \mathfrak{W}$ mögen *rangerniedrigend* heißen, wenn stets Rang $\mathfrak{W} \leq$ Rang $\mathfrak{Z}$ ist; sie mögen *rangerhaltend* heißen, wenn sogar stets Rang $\mathfrak{W} =$ Rang $\mathfrak{Z}$ ist[3]). Daß nicht jede lineare homogene Funktion rangerniedrigend ist, zeigt das Beispiel

$$\mathfrak{Z} = \begin{pmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{pmatrix}, \qquad f(\mathfrak{Z}) = \begin{pmatrix} z_{11} & 0 \\ 0 & z_{11} \end{pmatrix}.$$

Eine Funktion, die in der Gestalt $f(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}]$ mit von $\mathfrak{Z}$ unabhängiger Matrix $\mathfrak{A}$ geschrieben werden kann, heiße eine *spezielle* lineare homogene Funktion.

Satz 1. *Jede spezielle lineare homogene Funktion* $\mathfrak{Z}[\mathfrak{A}]$ *ist rangerniedrigend; sie ist rangerhaltend genau dann, wenn Rang* $\mathfrak{A} = n$ *ist.*

Beweis. Es ist Rang $\mathfrak{A}'\mathfrak{Z}\mathfrak{A} \leq$ Rang $\mathfrak{Z}$, woraus die Rangerniedrigung folgt. Der Rang von $\mathfrak{A}$ sei k. Dann ist $\mathfrak{A}$ in ein Produkt $\mathfrak{A} = \mathfrak{D}\mathfrak{C}\mathfrak{B}$ zerlegbar, wo $\mathfrak{B}$ eine m-reihige quadratische Matrix mit von Null verschiedener Determinante, $\mathfrak{D}$ eine n-reihige quadratische Matrix mit von Null verschiedener Determinante und $\mathfrak{C} = (c_{ij})$ eine Matrix ist, deren Elemente sämtlich verschwinden

[3]) „rangerniedrigend" und „Rangerniedrigung" schließen also stets die Möglichkeit ein, daß der Rang erhalten bleibt.

mit Ausnahme von $c_{11}=c_{22}=\cdots=c_{kk}=1$. Die Abbildung $\mathfrak{Z}\to\mathfrak{W}=\mathfrak{Z}[\mathfrak{A}]$ läßt sich dann in die Hintereinanderausführung der drei Abbildungen $\mathfrak{Z}\to\mathfrak{Z}_1=\mathfrak{Z}[\mathfrak{D}]$, $\mathfrak{Z}_1\to\mathfrak{Z}_2=\mathfrak{Z}_1[\mathfrak{C}]$, $\mathfrak{Z}_2\to\mathfrak{W}=\mathfrak{Z}_2[\mathfrak{B}]$ zerlegen. Die erste und dritte dieser Abbildungen sind umkehrbar und rangerhaltend und können als bloße Änderungen eines Koordinatensystems angesehen werden, die zweite Abbildung dagegen führt die n-reihige symmetrische Matrix

$$\begin{pmatrix} z_{11} & \cdots & z_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ z_{n1} & \cdots & z_{nn} \end{pmatrix}$$

in die m-reihige symmetrische Matrix

$$\begin{pmatrix} z_{11} & \cdots & z_{1k} & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ z_{k1} & \cdots & z_{kk} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

über; daraus lassen sich alle behaupteten Aussagen ablesen. — Über die in Satz 1 genannten Aussagen hinaus zeigt dieser Beweis, wie der Raum aller symmetrischen Matrizen $\mathfrak{Z}$ auf einen gewissen linearen Teilraum der symmetrischen Matrizen $\mathfrak{W}$ abgebildet wird.

Nicht jede rangerniedrigende lineare homogene Funktion ist speziell; Gegenbeispiel:

$$\mathfrak{Z}=\begin{pmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{pmatrix}, \qquad f(\mathfrak{Z})=\begin{pmatrix} z_{12} & 0 \\ 0 & 0 \end{pmatrix}.$$

Der folgende Satz zeigt jedoch, daß jede rangerhaltende lineare homogene Funktion speziell ist.

Satz 2. *Die lineare homogene Funktion $f(\mathfrak{Z})$ sei rangerniedrigend, und für mindestens eine feste Matrix $\mathfrak{Z}^{(0)}$ sei Rang $\mathfrak{Z}^{(0)}=$ Rang $f(\mathfrak{Z}^{(0)})=n$.* (Für rangerhaltende Funktionen sind diese Voraussetzungen gewiß erfüllt.) *Dann ist $f(\mathfrak{Z})$ rangerhaltend und speziell.*

Beweis. Man bestimme eine quadratische Matrix $\mathfrak{A}_0$ mit $\mathfrak{A}_0'\mathfrak{A}_0=\mathfrak{Z}^{(0)}$; dies ist möglich, und es ist $\det\mathfrak{A}_0\neq 0$. Man setze $f_1(\mathfrak{Z})=f(\mathfrak{Z}[\mathfrak{A}_0])$; dann ist $f_1(\mathfrak{Z})$ linear homogen und rangerniedrigend, und es ist Rang $f_1(\mathfrak{E})=n$. (In diesem Beweis soll $\mathfrak{E}$ immer die n-reihige Einheitsmatrix bedeuten.) $\mathfrak{Z}_k$ sei für $k=1,\ldots,n$ die Matrix, für die $z_{kk}=1$ ist und alle übrigen Elemente verschwinden. Da f_1 rangerniedrigend ist, hat $f_1(\mathfrak{Z}_k)$ höchstens den Rang 1. Wir bestimmen m-spaltige Zeilen $\mathfrak{a}_k$ mit $f_1(\mathfrak{Z}_k)=\mathfrak{a}_k'\mathfrak{a}_k$ und stellen

diese Zeilen zu einer n-zeiligen, m-spaltigen Matrix $\mathfrak{A}_1 = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_n \end{pmatrix}$ zusammen. Dann ist $f_1(\mathfrak{Z}_k) = \mathfrak{Z}_k[\mathfrak{A}_1]$ und folglich wegen der Linearität von f_1 auch $f_1(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}_1]$ für jede Diagonalmatrix $\mathfrak{Z}$. Durch Einsetzen von $\mathfrak{Z} = \mathfrak{E}$ ergibt sich $\operatorname{Rang} \mathfrak{A}_1' \mathfrak{A}_1 = n$, woraus $\operatorname{Rang} \mathfrak{A}_1 = n$ folgt. Wir ergänzen $\mathfrak{A}_1$ durch Hinzufügen von $m - n$ weiteren Zeilen zu einer quadratischen Matrix $\mathfrak{A}_2 = \begin{pmatrix} \mathfrak{A}_1 \\ * \end{pmatrix}$ dergestalt, daß $\mathfrak{A}_2$ eine Inverse besitzt. Dann ist $(\mathfrak{E}\ 0)\mathfrak{A}_2 = \mathfrak{A}_1$, also $\mathfrak{A}_1 \mathfrak{A}_2^{-1} = (\mathfrak{E}\ 0)$. Wir setzen $f_1(\mathfrak{Z})[\mathfrak{A}_2^{-1}] = f_2(\mathfrak{Z})$; diese Funktion ist wieder linear homogen und rangerniedrigend. Wir zerlegen für beliebiges $\mathfrak{Z}$

$$f_2(\mathfrak{Z}) = \begin{pmatrix} f_3(\mathfrak{Z}) & * \\ * & * \end{pmatrix},$$

wo $f_3(\mathfrak{Z})$ eine n-reihige quadratische Matrix bedeutet; auch die Funktion f_3 ist linear homogen und rangerniedrigend. Speziell für Diagonalmatrizen $\mathfrak{Z}$ ist dann

$$f_2(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}_1 \mathfrak{A}_2^{-1}] = \mathfrak{Z}[(\mathfrak{E}\,0)] = \begin{pmatrix} \mathfrak{Z} & 0 \\ 0 & 0 \end{pmatrix},$$

also $f_3(\mathfrak{Z}) = \mathfrak{Z}$. Für beliebige Matrizen $\mathfrak{Z}$ gilt wegen der Linearität der Funktion f_3

$$f_3(\lambda\mathfrak{E} - \mathfrak{Z}) = \lambda\mathfrak{E} - f_3(\mathfrak{Z}),$$

und da f_3 rangerniedrigend ist, folgt für festes λ aus $\det(\lambda\mathfrak{E} - \mathfrak{Z}) = 0$ die Gleichung $\det(\lambda\mathfrak{E} - f_3(\mathfrak{Z})) = 0$. Falls die Eigenwerte von $\mathfrak{Z}$ paarweise verschieden sind, haben also $\det(\lambda\mathfrak{E} - \mathfrak{Z})$ und $\det(\lambda\mathfrak{E} - f_3(\mathfrak{Z}))$ als Polynome in λ die gleichen Nullstellen und die gleichen höchsten Koeffizienten. Die daraus folgende Gleichung $\det(\lambda\mathfrak{E} - \mathfrak{Z}) = \det(\lambda\mathfrak{E} - f_3(\mathfrak{Z}))$ ist eine Identität in λ und den Elementen von $\mathfrak{Z}$ und gilt deshalb ohne Einschränkung. Insbesondere ist $\det \mathfrak{Z} = \det f_3(\mathfrak{Z})$. Man setze $\mathfrak{Z}_{ik}$ gleich der Matrix, in der $z_{ik} = z_{ki} = 1$ ist und alle übrigen Elemente verschwinden, und verstehe unter

$$\mathfrak{Z}_0 = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix}$$

eine Diagonalmatrix mit unbestimmten Elementen. Aus

$$\det(\mathfrak{Z}_0 + \mathfrak{Z}_{ik}) = \det f_3(\mathfrak{Z}_0 + \mathfrak{Z}_{ik}) = \det(\mathfrak{Z}_0 + f_3(\mathfrak{Z}_{ik}))$$

folgt durch Vergleich der Koeffizienten der Produkte der $z_1, \ldots, z_n$, daß die symmetrische Matrix $f_3(\mathfrak{Z}_{ik})$ gleich $\pm\mathfrak{Z}_{ik}$ ist. $\mathfrak{A}_3$ sei eine Diagonalmatrix mit Diagonalelementen ± 1. Man setze $f_4(\mathfrak{Z}) = f_3(\mathfrak{Z})[\mathfrak{A}_3]$. Diese Funktion f_4 besitzt alle bei der Funktion f_3 konstatierten Eigenschaften, und durch eine stets mögliche passende Wahl der Vorzeichen in der Diagonale von $\mathfrak{A}_3$ sei für $2 \leq k \leq n$ ferner erreicht, daß $f_4(\mathfrak{Z}_{1k}) = +\mathfrak{Z}_{1k}$ ist. Sei nun $2 \leq i, k \leq n$ und $i \neq k$. Vergleicht man in

$$\begin{aligned} \det(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + \mathfrak{Z}_{ik}) &= \det f_4(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + \mathfrak{Z}_{ik}) \\ &= \det(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + f_4(\mathfrak{Z}_{ik})) \end{aligned}$$

wieder die Koeffizienten der Produkte der $z_1, \ldots, z_n$, so erkennt man, daß jetzt $f_4(\mathfrak{Z}_{ik}) = +\mathfrak{Z}_{ik}$ ist; wegen der Linearität von f_4 gilt also $f_4(\mathfrak{Z}) = \mathfrak{Z}$ für alle $\mathfrak{Z}$. Wir bezeichnen mit $\mathfrak{E}^*$ die $(m-n)$-reihige Einheitsmatrix, setzen $\mathfrak{A}_4 = \begin{pmatrix} \mathfrak{A}_3 & 0 \\ 0 & \mathfrak{E}^* \end{pmatrix}$ und betrachten die lineare homogene Funktion

$$f_5(\mathfrak{Z}) = f_2(\mathfrak{Z})\,[\mathfrak{A}_4] = \begin{pmatrix} \mathfrak{Z} & f_6(\mathfrak{Z}) \\ (f_6(\mathfrak{Z}))' & f_7(\mathfrak{Z}) \end{pmatrix}. \tag{10}$$

Weil $\det \mathfrak{A}_4 \neq 0$ und $f_2(\mathfrak{Z})$ rangerniedrigend ist, ist Rang $f_5(\mathfrak{Z}) \leq$ Rang $\mathfrak{Z}$; andererseits folgt aus der Gestalt der rechten Seite von (10) Rang $f_5(\mathfrak{Z}) \geq$ Rang $\mathfrak{Z}$. Die Funktion $f_5(\mathfrak{Z})$ ist demnach rangerhaltend; also sind auch $f_2(\mathfrak{Z})$, $f_1(\mathfrak{Z})$ und schließlich auch das ursprüngliche $f(\mathfrak{Z})$ rangerhaltend. Für Diagonalmatrizen $\mathfrak{Z}$ verschwinden $f_6(\mathfrak{Z})$ und $f_7(\mathfrak{Z})$. Wir zeigen, daß $f_6(\mathfrak{Z})$ identisch verschwindet; dazu genügt wegen der Linearität von f_6 der Nachweis, daß $f_6(\mathfrak{Z}_{ik})$ für $1 \leq i, k \leq n$ und $i \neq k$ verschwindet. Die Matrizen $\mathfrak{Z}^{(1)} = \mathfrak{Z}_i + \mathfrak{Z}_k + \mathfrak{Z}_{ik}$ und $\mathfrak{Z}^{(2)} = -\mathfrak{Z}_i - \mathfrak{Z}_k + \mathfrak{Z}_{ik}$ haben den Rang 1; es ist also für $j = 1, 2$ auch

$$\text{Rang } f_5(\mathfrak{Z}^{(j)}) = \text{Rang} \begin{pmatrix} \mathfrak{Z}^{(j)} & f_6(\mathfrak{Z}_{ik}) \\ * & * \end{pmatrix} = 1.$$

Die Spalten von $f_6(\mathfrak{Z}_{ik})$ müssen also sowohl zu den Spalten von $\mathfrak{Z}^{(1)}$ als auch zu den Spalten von $\mathfrak{Z}^{(2)}$ proportional sein, woraus ihr Verschwinden folgt. Also ist für alle $\mathfrak{Z}$

$$f_5(\mathfrak{Z}) = \begin{pmatrix} \mathfrak{Z} & 0 \\ 0 & f_7(\mathfrak{Z}) \end{pmatrix};$$

da f_5 rangerniedrigend ist, folgt daraus $f_7(\mathfrak{Z}) = 0$. Wir haben nun

$$f(\mathfrak{Z}\,[\mathfrak{A}_0])\,[\mathfrak{A}_2^{-1}\,\mathfrak{A}_4] = f_5(\mathfrak{Z}) = \mathfrak{Z}\,[(\mathfrak{E}\,0)],$$

woraus $f(\mathfrak{Z}) = \mathfrak{Z}\,[\mathfrak{A}]$ mit $\mathfrak{A} = \mathfrak{A}_0^{-1}(\mathfrak{E}\,0)\,\mathfrak{A}_4^{-1}\,\mathfrak{A}_2$ folgt. $f(\mathfrak{Z})$ ist also speziell.

Satz 3. *Ist $f(\mathfrak{Z}) = \mathfrak{Z}\,[\mathfrak{A}]$ eine spezielle lineare homogene Funktion, so ist durch $f(\mathfrak{Z})$ die Matrix $\mathfrak{A}$ bis auf einen willkürlich bleibenden skalaren Faktor ± 1 eindeutig bestimmt.*

Beweis. Aus $\mathfrak{A}'\mathfrak{Z}\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}\mathfrak{B}$ für alle $\mathfrak{Z}$ haben wir $\mathfrak{A} = \pm\mathfrak{B}$ zu schließen. $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ seien die Zeilen von $\mathfrak{A}$, und $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ seien die Zeilen von $\mathfrak{B}$, ferner mögen $\mathfrak{Z}_k$ und $\mathfrak{Z}_{ik}$ dieselbe Bedeutung wie im Beweis von Satz 2 haben. $\mathfrak{A}'\mathfrak{Z}_k\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}_k\mathfrak{B}$ ergibt $\mathfrak{a}_k'\mathfrak{a}_k = \mathfrak{b}_k'\mathfrak{b}_k$, also $\mathfrak{a}_k = \pm\mathfrak{b}_k$. Angenommen, in dieser Gleichung gälte nicht für alle Indizes k das gleiche Vorzeichen; etwa für die Indizes i, k sei $\mathfrak{a}_i = +\mathfrak{b}_i$ und $\mathfrak{a}_k = -\mathfrak{b}_k$; überdies sei $\mathfrak{a}_i \neq 0$ und $\mathfrak{a}_k \neq 0$ (da sonst das betreffende Vorzeichen sowieso irrelevant wäre). Aus $\mathfrak{A}'\mathfrak{Z}_{ik}\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}_{ik}\mathfrak{B}$ folgte dann $\mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = \mathfrak{b}_i'\mathfrak{b}_k + \mathfrak{b}_k'\mathfrak{b}_i = \mathfrak{a}_i'(-\mathfrak{a}_k) + (-\mathfrak{a}_k)'\mathfrak{a}_i$, also $\mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = 0$. Eine solche Gleichung kann nur bestehen, wenn die Zeilen $\mathfrak{a}_i$ und $\mathfrak{a}_k$ proportional sind, also etwa $\mathfrak{a}_i = \lambda\mathfrak{a}_k$. Das gibt $2\lambda\,\mathfrak{a}_k'\mathfrak{a}_k = \mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = 0$. Da die Matrix $\mathfrak{a}_k'\mathfrak{a}_k$ nicht die Nullmatrix ist, folgt $\lambda = 0$, $\mathfrak{a}_i = 0$ entgegen der Annahme.

§ 3. Differentiell rangerniedrigende und differentiell spezielle Funktionen.

Wir betrachten nun beliebige differenzierbare Funktionen $\mathfrak{W} = f(\mathfrak{Z})$; darunter soll verstanden werden, daß die Elemente von $\mathfrak{W}$ total differenzierbare Funktionen der Elemente von $\mathfrak{Z}$ sind. Die Matrix der Differentiale $d\mathfrak{W}$ ist dann eine lineare homogene Funktion der Matrix der Differentiale $d\mathfrak{Z}$. Wir nennen $\mathfrak{W} = f(\mathfrak{Z})$ eine *differentiell rangerniedrigende* bzw. *differentiell spezielle* bzw. *differentiell rangerhaltende* Funktion, wenn $d\mathfrak{W}$ für alle in Betracht kommenden $\mathfrak{Z}$ eine rangerniedrigende bzw. spezielle bzw. rangerhaltende Funktion von $d\mathfrak{Z}$ ist. Die Hintereinanderausführung differentiell rangerniedrigender bzw. differentiell spezieller bzw. differentiell rangerhaltender Funktionen ergibt, wenn sie ausführbar ist, wieder eine solche Funktion; im Falle $m = n$ ist für differentiell rangerhaltende Funktionen eine Umkehrfunktion vorhanden, und diese ist wieder differentiell rangerhaltend. — Für beliebige differentiell spezielle Funktionen könnte man die in der Relation

$$d\mathfrak{W} = (d\mathfrak{Z})\,[\mathfrak{M}] \tag{11}$$

auftretende, bis auf ihr Vorzeichen bestimmte und im allgemeinen von $\mathfrak{Z}$ abhängende Matrix $\mathfrak{M}$ die symmetrische Ableitung von $f(\mathfrak{Z})$ nennen.

Wir wollen uns noch durch einige Beispiele überzeugen, daß es differentiell spezielle Funktionen gibt[4]). Ein triviales Beispiel solcher Funktionen sind die speziellen linearen homogenen Funktionen; ein weiteres wichtiges Beispiel wird durch die symplektische gebrochen lineare Funktion

$$\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})\,(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$$

geliefert, für diese gilt nämlich $d\mathfrak{W} = (d\mathfrak{Z})\,[(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}]$. Die Hintereinanderausführung der in diesen beiden Beispielen auftretenden Funktionen liefert eine Klasse von Funktionen, die in der Elektrotechnik bei der Transformation von symmetrischen Widerstandsmatrizen auftreten; man vergleiche die Bemerkungen am Ende von § 1.

Der Rest dieses Paragraphen bezieht sich auf den Spezialfall $m = 1$. Die Matrix $\mathfrak{W}$ besteht dann nur aus einem Element, das eine Funktion $w = w(\mathfrak{Z})$ ist; die Matrix $\mathfrak{M} = (\mu_1, \ldots, \mu_n)$ besteht nur aus einer Zeile, und die $\mu_1, \ldots, \mu_n$ sind im allgemeinen ebenfalls Funktionen von $\mathfrak{Z}$. Nach (11) sind die Gleichungen

$$\frac{\partial w}{\partial z_{ik}} = \varepsilon_{ik}\,\mu_i\,\mu_k \qquad (i = 1, \ldots, n;\ k = 1, \ldots, n) \tag{12}$$

zu erfüllen, abkürzend ist

$$\varepsilon_{ik} = \begin{cases} 1 & \text{für} \quad i = k, \\ 2 & \text{für} \quad i \neq k \end{cases}$$

[4]) Die Frage nach der Aufstellung aller dieser Funktionen, d. h. nach der Bestimmung aller Lösungen des Systems partieller Differentialgleichungen (11), bleibt offen.

gesetzt. Aus (12) läßt sich leicht ein System von Differentialgleichungen herleiten, in denen $\mu_1, \ldots, \mu_n$ nicht mehr vorkommen:

$$\frac{1}{\varepsilon_{ik}\varepsilon_{jl}} \frac{\partial w}{\partial z_{ik}} \frac{\partial w}{\partial z_{jl}} = \frac{1}{\varepsilon_{ij}\varepsilon_{kl}} \frac{\partial w}{\partial z_{ij}} \frac{\partial w}{\partial z_{kl}}, \tag{13}$$

sowie ein System von Integrabilitätsbedingungen:

$$\varepsilon_{ik} \frac{\partial(\mu_i \mu_k)}{\partial z_{jl}} = \varepsilon_{jl} \frac{\partial(\mu_j \mu_l)}{\partial z_{ik}}$$ [5]).

Da einige Lösungen der Gln. (12) bzw. (13) bekannt sind, die einige Parameter enthalten (z. B. $w = \mathfrak{Z}[\mathfrak{M}]$ mit konstantem $\mathfrak{M} = (\mu_1, \ldots, \mu_n)$, wo $\mu_1, \ldots, \mu_n$ als Parameter aufgefaßt werden mögen), so kann man aus diesen in üblicher Weise durch Differentiations- und Eliminationsprozesse weitere Lösungen des Gleichungssystems (13) gewinnen. Man kann zur Auffindung weiterer Lösungen von (12), (13) auch folgendermaßen vorgehen: $\varphi_{ik}(w)$ $(i = 1, \ldots, n;\ k = 1, \ldots, n;\ \varphi_{ik}(w) = \varphi_{ki}(w))$ seien irgendwie vorgegebene differenzierbare Funktionen von w; es sei möglich, aus der Gleichung

$$\det\left(z_{ik} - \varphi_{ik}(w)\right) = 0 \tag{14}$$

die Größe w als Funktion der Elemente von $\mathfrak{Z}$ zu bestimmen. $\nu_1, \ldots, \nu_n$ sei eine nichttriviale Lösung des Systems

$$\sum_{i=1}^{n} \left(z_{ik} - \varphi_{ik}(w)\right) \nu_i = 0.$$

[Für w ist hier und im folgenden immer die aus (14) bestimmte Funktion von $\mathfrak{Z}$ einzusetzen; $\nu_1, \ldots, \nu_n$ hängen also ebenfalls von $\mathfrak{Z}$ ab.] $\varphi'_{ik}(w)$ sei die Ableitung von $\varphi_{ik}(w)$. Der Ausdruck

$$\sum_{i,k} \varphi'_{ik}(w)\, \nu_i \nu_k$$

sei von Null verschieden; es werde für $j = 1, \ldots, n$

$$\mu_j = \mu_j(\mathfrak{Z}) = \frac{\nu_j}{\sqrt{\sum_{i,k} \varphi'_{ik}(w)\, \nu_i \nu_k}}$$

gesetzt. Dann gilt

$$\sum_{i,k} \varphi'_{ik}(w)\, \mu_i \mu_k = 1 \tag{15}$$

und

$$\sum_{i} \left(z_{ik} - \varphi_{ik}(w)\right) \mu_i = 0 \qquad (k = 1, \ldots, n). \tag{16}$$

[5]) Bei Einführung neuer Variabeln $\zeta_{ik} = \varepsilon_{ik} z_{ik}$ vereinfachen sich (12) und die daraus hergeleiteten Gleichungen noch etwas durch den Wegfall der ε-Faktoren. — Für $m > 1$ lassen sich entsprechende Gleichungen aufstellen.

Differentiation der daraus folgenden Gleichung

$$\sum_{i,k} \left(z_{ik} - \varphi_{ik}(w)\right) \mu_i \mu_k = 0$$

liefert

$$\sum_{i,k} \left(dz_{ik} - \varphi'_{ik}(w)\, dw\right) \mu_i \mu_k + 2 \sum_{i,k} \left(z_{ik} - \varphi_{ik}(w)\right) \mu_i\, d\mu_k = 0,$$

also unter Benutzung von (15) und (16)

$$\sum_{i,k} dz_{ik}\, \mu_i \mu_k - dw = 0;$$

das ist (12) in anderer Schreibweise.

Für $m=1$, $n=2$ reduziert sich (13) auf

$$4 \frac{\partial w}{\partial z_{11}} \cdot \frac{\partial w}{\partial z_{22}} - \left(\frac{\partial w}{\partial z_{12}}\right)^2 = 0.$$

Im euklidischen (z_{11}, z_{12}, z_{22})-Raum gedeutet sind die Flächen $w=\text{const}$ also Torsen, deren Tangentialebenen parallel zu den Tangentialebenen eines festen Kegels verlaufen. Eine Lösung ist beispielsweise $w = 2z_{12} + 2\sqrt{z_{11} z_{22}}$; für dieses w ist

$$dw = \begin{pmatrix} dz_{11} & dz_{12} \\ dz_{12} & dz_{22} \end{pmatrix} \begin{bmatrix} \sqrt[4]{\frac{z_{22}}{z_{11}}} \\ \sqrt[4]{\frac{z_{11}}{z_{22}}} \end{bmatrix}.$$

§ 4. Die Eigenwerte bei rangerniedrigender linearer Abbildung

Jeder m-reihigen komplexen quadratischen Matrix $\mathfrak{M} = \mathfrak{M}^{(1)} + i\mathfrak{M}^{(2)}$ werde die $2m$-reihige reelle quadratische Matrix

$$\widetilde{\mathfrak{M}} = \begin{pmatrix} \mathfrak{M}^{(1)} & \mathfrak{M}^{(2)} \\ \mathfrak{M}^{(2)} & -\mathfrak{M}^{(1)} \end{pmatrix}$$

zugeordnet und umgekehrt. Ist $\mathfrak{Z}' = \mathfrak{Z}$, so ist offenbar $\widetilde{\mathfrak{Z}}' = \widetilde{\mathfrak{Z}}$ und umgekehrt. Ist $\mathfrak{U}$ unitär, so ist $\widetilde{\mathfrak{U}}$ orthogonal und umgekehrt, wie man sogleich erkennt, wenn man die Unitaritätsbedingung $\mathfrak{U}'\overline{\mathfrak{U}} = \mathfrak{E}$ (m-reihige Einheitsmatrix) in Real- und Imaginärteil spaltet und dann mit der Orthogonalitätsbedingung $\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{U}} = \mathfrak{E}$ ($2m$-reihige Einheitsmatrix) vergleicht.

Hilfssatz. *Sei* $\mathfrak{W}' = \mathfrak{W}$, *und* $\mathfrak{W}\overline{\mathfrak{W}}$ *habe die (eo ipso reellen und nichtnegativen) Eigenwerte* $q_1^2 \geq \cdots \geq q_m^2$; *dabei sei* $q_\mu \geq 0$ *für* $\mu = 1, \ldots, m$. *Dann hat* $\widetilde{\mathfrak{W}}$ *die Eigenwerte* $q_1 \geq \cdots \geq q_m \geq -q_m \geq \cdots \geq -q_1$, *und* $\widetilde{\mathfrak{W}}^2$ *hat die Eigenwerte* $q_1^2 = q_1^2 \geq q_2^2 = q_2^2 \geq \cdots \geq q_m^2 = q_m^2$.

Beweis. Man setze

$$\mathfrak{Q} = \begin{pmatrix} q_1 & & \\ & \ddots & \\ & & q_m \end{pmatrix}.$$

Nach SIEGEL [2], p. 14–15 gibt es eine unitäre Matrix $\mathfrak{U}$ so, daß $\mathfrak{W}=\mathfrak{Q}[\mathfrak{U}]$ ist. Eine einfache Rechnung zeigt dann $\widetilde{\mathfrak{W}}=\widetilde{\mathfrak{Q}}[\widetilde{\mathfrak{U}}]$. Dabei ist $\widetilde{\mathfrak{U}}$ orthogonal; es folgt $\widetilde{\mathfrak{W}}^2=\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{Q}}\,\widetilde{\mathfrak{U}}\,\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{Q}}\,\widetilde{\mathfrak{U}}=\widetilde{\mathfrak{Q}}^2[\widetilde{\mathfrak{U}}]$. Wegen

$$\widetilde{\mathfrak{Q}}=\begin{pmatrix}\mathfrak{Q} & \\ & -\mathfrak{Q}\end{pmatrix} \quad \text{und} \quad \widetilde{\mathfrak{Q}}^2=\begin{pmatrix}\mathfrak{Q}^2 & \\ & \mathfrak{Q}^2\end{pmatrix}$$

folgen die Behauptungen.

Korollar 1. *Ist* $\mathfrak{E}-\mathfrak{W}\overline{\mathfrak{W}}>0$, *so ist* $\mathfrak{E}-\widetilde{\mathfrak{W}}^2>0$.

Korollar 2. *Ist* $\mathfrak{W}'=\mathfrak{W}$ *und Rang* $\mathfrak{W}\leq 1$, *so ist Rang* $\widetilde{\mathfrak{W}}\leq 2$.

Hauptsatz. *Die Abbildung* $\mathfrak{Z}\to\mathfrak{W}$ *sei linear homogen und rangerniedrigend, und für* $\mathfrak{E}-\mathfrak{Z}\overline{\mathfrak{Z}}>0$ *sei* $\mathfrak{E}-\mathfrak{W}\overline{\mathfrak{W}}>0$. *Die Eigenwerte von* $\mathfrak{Z}\overline{\mathfrak{Z}}$ *seien* $p_1^2\geq\cdots\geq p_n^2$, *die Eigenwerte von* $\mathfrak{W}\overline{\mathfrak{W}}$ *seien* $q_1^2\geq\cdots\geq q_m^2$. *Dann ist* $q_k^2\leq p_k^2$ *für* $1\leq k\leq\min(m,n)$ *und (falls* $n<m$ *ist)* $q_k^2=0$ *für* $n<k\leq m$.

Beweis. Für eine r-reihige reelle symmetrische Matrix $\mathfrak{M}$ lassen sich die nach abnehmender Größe geordneten Eigenwerte bekanntlich ([3], [4]) wie folgt bestimmen: $\mathfrak{x}$ durchlaufe alle Vektoren eines mindestens $(r-k+1)$-dimensionalen Teilraums T_k des r-dimensionalen reellen Vektorraums mit $\mathfrak{x}'\mathfrak{x}=1$. Man bestimme dafür $\operatorname{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{M}[\mathfrak{x}]$; dies Maximum hängt noch von T_k ab. Man variiere nun T_k und bestimme das Minimum dieser Maxima. Die so gefundene Zahl $\operatorname{Min}_{T_k}\operatorname{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{M}[\mathfrak{x}]$ ist dann der k-te Eigenwert von $\mathfrak{M}$.

Wir wenden dies Verfahren auf die $2m$-reihige Matrix $\widetilde{\mathfrak{W}}^2=\widetilde{\mathfrak{W}}'\widetilde{\mathfrak{W}}$ an. Der $2k$-te Eigenwert dieser Matrix ist nach dem Hilfssatz gleich q_k^2; also ist

$$q_k^2=\operatorname{Min}_{T_{2k}}\operatorname{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{y}'\mathfrak{y}, \tag{17}$$

wobei $\mathfrak{y}=\widetilde{\mathfrak{W}}\mathfrak{x}$ gesetzt wurde. Nach dem bereits genannten Satz von SIEGEL gilt mit $\mathfrak{P}=\begin{pmatrix}p_1 & & \\ & \ddots & \\ & & p_n\end{pmatrix}$ und einer passenden unitären Matrix $\mathfrak{U}$ die Gleichung $\mathfrak{Z}=\mathfrak{P}[\mathfrak{U}]$. Es werde $\mathfrak{Z}_k=\begin{pmatrix}0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & 0\end{pmatrix}[\mathfrak{U}]$ gesetzt, wobei in der hingeschriebenen Diagonalmatrix das k-te Diagonalelement gleich 1 ist und alle übrigen verschwinden. Das Bild von $\mathfrak{Z}_k$ sei $\mathfrak{W}_k$; wegen der Rangerniedrigung ist $\operatorname{Rang}\mathfrak{W}_k\leq\operatorname{Rang}\mathfrak{Z}_k=1$, und wegen der Linearität der Abbildung $\mathfrak{Z}\to\mathfrak{W}$ ist $\mathfrak{W}=p_1\mathfrak{W}_1+\cdots+p_n\mathfrak{W}_n$. Da $p_1,\ldots,p_n$ reell sind, folgt daraus

$$\widetilde{\mathfrak{W}}=p_1\widetilde{\mathfrak{W}}_1+\cdots+p_n\widetilde{\mathfrak{W}}_n.$$

Das obige $\mathfrak{y}$ ist also durch

$$\mathfrak{y}=(p_1\widetilde{\mathfrak{W}}_1+\cdots+p_n\widetilde{\mathfrak{W}}_n)\,\mathfrak{x} \tag{18}$$

bestimmt. Sei $1 \leq k \leq \min(m, n)$. Die linearen Gleichungen

$$\mathfrak{W}_1 \mathfrak{x} = 0, \quad \mathfrak{W}_2 \mathfrak{x} = 0, \ldots, \mathfrak{W}_{k-1} \mathfrak{x} = 0 \tag{19}$$

definieren einen mindestens $(2m - 2k + 2)$-dimensionalen Teilraum $T^{(0)}$ des $2n$-dimensionalen Vektorraums, denn nach Korollar 2 ist Rang $\mathfrak{W}_\nu \leq 2$, so daß also das System (19) höchstens $2(k-1)$ linear unabhängige Gleichungen enthält. Wir verkleinern nun die rechte Seite von (17) nicht, wenn wir an Stelle der Minimumbildung über alle T_{2k} nur den speziellen Raum $T^{(0)}$ heranziehen. Also ist

$$q_k^2 \leq \operatorname{Max} \mathfrak{y}' \mathfrak{y}, \tag{20}$$

wobei $\mathfrak{y}$ durch (18) gegeben ist und $\mathfrak{x}$ alle Vektoren zu durchlaufen hat, die den Bedingungen (19) und der Normierungsbedingung $\mathfrak{x}' \mathfrak{x} = 1$ genügen. Für diese speziellen $\mathfrak{x}$ folgt aus (18) und (19)

$$\mathfrak{y} = (p_k \mathfrak{W}_k + \cdots + p_n \mathfrak{W}_n)\, \mathfrak{x}.$$

Entweder ist nun $p_k = p_{k+1} = \cdots = p_n = 0$; dann ist $\mathfrak{y} = 0$ und gewiß $\mathfrak{y}' \mathfrak{y} \leq p_k^2$. Oder es ist $p_k > 0$; in diesem Falle schreiben wir $\mathfrak{y} = p_k \mathfrak{M} \mathfrak{x}$ mit

$$\mathfrak{M} = \frac{p_k}{p_k} \mathfrak{W}_k + \frac{p_{k+1}}{p_k} \mathfrak{W}_{k+1} + \cdots + \frac{p_n}{p_k} \mathfrak{W}_n.$$

$\mathfrak{M}$ ist dabei das Bild der Matrix

$$\mathfrak{Z}^{(0)} = \frac{p_k}{p_k} \mathfrak{Z}_k + \frac{p_{k+1}}{p_k} \mathfrak{Z}_{k+1} + \cdots + \frac{p_n}{p_k} \mathfrak{Z}_n,$$

und $\mathfrak{Z}^{(0)} \overline{\mathfrak{Z}^{(0)}}$ hat die Eigenwerte $0, \ldots, 0, \left(\frac{p_k}{p_k}\right)^2, \left(\frac{p_{k+1}}{p_k}\right)^2, \ldots, \left(\frac{p_n}{p_k}\right)^2$, die sämtlich zwischen 0 und 1 liegen. Also ist $\mathfrak{E} - \mathfrak{Z}^{(0)} \overline{\mathfrak{Z}^{(0)}} \geq 0$, also $\mathfrak{E} - \mathfrak{M} \overline{\mathfrak{M}} \geq 0$, also nach Korollar 1 $\mathfrak{E} - \mathfrak{M}^2 \geq 0$, also, da $\mathfrak{x}' \mathfrak{x} = 1$ ist, $\mathfrak{x}' \mathfrak{M}' \mathfrak{M} \mathfrak{x} = \mathfrak{M}^2[\mathfrak{x}] \leq 1$, also $\mathfrak{y}' \mathfrak{y} \leq p_k^2$. In jedem Fall gilt also $\mathfrak{y}' \mathfrak{y} \leq p_k^2$, und aus (20) folgt die Behauptung $q_k^2 \leq p_k^2$. — Die für den Fall $n < m$ gemachte zusätzliche Behauptung ist trivial, da wegen der vorausgesetzten Rangerniedrigung stets Rang $\mathfrak{W} \leq$ Rang $\mathfrak{Z} \leq n$ ist.

Anmerkung. Ohne die Voraussetzung der Rangerniedrigung ist der vorstehende Satz falsch, wie etwa das Beispiel $\mathfrak{Z} \to \mathfrak{W} = \begin{pmatrix} \mathfrak{Z} & \\ & \mathfrak{Z} \end{pmatrix}$ zeigt.

§ 5. Ein Analogon zum SCHWARZschen Lemma

In diesem Paragraphen sei $\mathfrak{W} = \varphi(\mathfrak{Z})$ eine differentiell rangerniedrigende analytische Matrixfunktion, die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes abbildet.

Satz 1 (SIEGEL). *Wird bei $\mathfrak{Z} \to \mathfrak{W}$ der Nullpunkt auf den Nullpunkt abgebildet, so gilt dort $d\mathfrak{W} = f(d\mathfrak{Z})$, wobei f eine lineare homogene Funktion ist, die den Einheitskreis in den Einheitskreis abbildet.*

Beweis. $\mathfrak{W}_1$ sei der lineare Bestandteil der Entwicklung der Elemente von $\mathfrak{W}$ nach Potenzprodukten der Elemente von $\mathfrak{Z}$; also ist $\mathfrak{W}_1 = f(\mathfrak{Z})$ eine lineare homogene Funktion von $\mathfrak{Z}$. An der Stelle $\mathfrak{Z} = 0$ ist dann $d\mathfrak{W} = f(d\mathfrak{Z})$. Es bleibt zu zeigen, daß für $\mathfrak{E} - \mathfrak{Z}\overline{\mathfrak{Z}} > 0$ auch $\mathfrak{E} - \mathfrak{W}_1\overline{\mathfrak{W}}_1 > 0$ ist; für den Beweis dieser Tatsache sei auf SIEGEL [2], p. 11–12 verwiesen. — Bisher wurde die differentielle Rangerniedrigung von φ nicht benötigt; sie liefert uns jedoch unmittelbar den

Zusatz. Das obengenannte f ist eine rangerniedrigende lineare homogene Funktion.

Satz 2. *ds sei ein symplektisches Bogenelement im Einheitskreis des $\mathfrak{Z}$-Raumes, dt sei sein Bild im $\mathfrak{W}$-Raum. Dann ist $dt \leq ds$.*

Beweis. Das Bogenelement ds liege an der Stelle $\mathfrak{Z} = \mathfrak{Z}_0$, und das Bogenelement dt liege an der Stelle $\mathfrak{W} = \mathfrak{W}_0$. Durch passende symplektische linear gebrochene Funktionen, die den Einheitskreis des $\mathfrak{Z}$-Raumes bzw. den Einheitskreis des $\mathfrak{W}$-Raumes auf sich abbilden und alle symplektischen Bogenlängen ungeändert lassen, können $\mathfrak{Z}_0$ und $\mathfrak{W}_0$ in die Nullmatrizen übergeführt werden (vgl. [2]). Diese symplektischen linear gebrochenen Funktionen sind nach § 3 differentiell rangerhaltend und ergeben bei Zusammensetzung mit φ eine differentiell rangerniedrigende Funktion, die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes abbildet und den Nullpunkt in den Nullpunkt überführt. Man kann daher ohne Beschränkung der Allgemeinheit $\mathfrak{Z}_0 = 0$ und $\mathfrak{W}_0 = 0$ annehmen. Nach (8) ist im Nullpunkt $ds^2 = 4 \cdot \text{Spur}\,(d\mathfrak{Z}\,d\overline{\mathfrak{Z}})$ und entsprechend $dt^2 = 4 \cdot \text{Spur}\,(d\mathfrak{W}\,d\overline{\mathfrak{W}})$. Die Funktion $\mathfrak{W}_1 = f(\mathfrak{Z})$ erfüllt nach Satz 1 die Voraussetzungen des Hauptsatzes von § 4. Die nach abnehmender Größe geordneten Eigenwerte von $\mathfrak{W}_1\overline{\mathfrak{W}}_1$ sind daher jeweils nicht größer als die entsprechenden nach abnehmender Größe geordneten Eigenwerte von $\mathfrak{Z}\overline{\mathfrak{Z}}$. Wegen Satz 1 gilt dasselbe für den Vergleich der Eigenwerte von $d\mathfrak{W}\,d\overline{\mathfrak{W}}$ und $d\mathfrak{Z}\,d\overline{\mathfrak{Z}}$. Folglich ist $dt^2 \leq ds^2$ wie behauptet.

Aus Satz 2 folgt durch Integration

Satz 3. *Bei der Abbildung $\mathfrak{Z} \to \mathfrak{W} = \varphi(\mathfrak{Z})$ nimmt die symplektische Bogenlänge jeder im Einheitskreis gelegenen Kurve nicht zu.*

Satz 4. *$\mathfrak{Z}^{(1)}$ und $\mathfrak{Z}^{(2)}$ seien zwei Punkte im Einheitskreis; $\mathfrak{W}^{(1)} = \varphi(\mathfrak{Z}^{(1)})$ und $\mathfrak{W}^{(2)} = \varphi(\mathfrak{Z}^{(2)})$ seien ihre Bilder. Dann ist die symplektische Entfernung von $\mathfrak{W}^{(1)}$, $\mathfrak{W}^{(2)}$ höchstens gleich der symplektischen Entfernung von $\mathfrak{Z}^{(1)}$, $\mathfrak{Z}^{(2)}$.*

Beweis. Man verbinde $\mathfrak{Z}^{(1)}$ und $\mathfrak{Z}^{(2)}$ durch einen geodätischen Bogen; dessen symplektische Länge l ist dann gleich der symplektischen Entfernung von $\mathfrak{Z}^{(1)}$, $\mathfrak{Z}^{(2)}$ (vgl. [2]). Bei der Abbildung $\mathfrak{Z} \to \mathfrak{W}$ geht dieser Bogen in eine Kurve über, deren Bogenlänge l' nach Satz 3 höchstens gleich l ist. Die symplektische Entfernung l'' von $\mathfrak{W}^{(1)}$, $\mathfrak{W}^{(2)}$ ist höchstens gleich der Länge l' des letztgenannten Bogens, der ja $\mathfrak{W}^{(1)}$ und $\mathfrak{W}^{(2)}$ verbindet. Also ist $l'' \leq l$ wie behauptet.

Die Frage bleibt offen, ob bzw. unter welchen Bedingungen für die nach abnehmender Größe geordneten Eigenwerte der gemäß (7) gebildeten Matrizen

$$(\mathfrak{Z}^{(1)} - \mathfrak{Z}^{(2)})\,(\mathfrak{E} - \overline{\mathfrak{Z}}^{(2)}\,\mathfrak{Z}^{(1)})^{-1}\,(\overline{\mathfrak{Z}}^{(1)} - \overline{\mathfrak{Z}}^{(2)})\,(\mathfrak{E} - \mathfrak{Z}^{(2)}\,\overline{\mathfrak{Z}}^{(1)})^{-1}$$

und

$$(\mathfrak{W}^{(1)} - \mathfrak{W}^{(2)})\,(\mathfrak{E} - \overline{\mathfrak{W}}^{(2)}\,\mathfrak{W}^{(1)})^{-1}\,(\overline{\mathfrak{W}}^{(1)} - \overline{\mathfrak{W}}^{(2)})\,(\mathfrak{E} - \mathfrak{W}^{(2)}\,\overline{\mathfrak{W}}^{(1)})^{-1}$$

entsprechende Ungleichungen gelten wie im Hauptsatz von § 4.

Literatur

[1] Fuhrmann, A.: Klassen ähnlicher Matrizen als verallgemeinerte Doppelverhältnisse. Math. Z. **62**, 211—240 (1955). — [2] Siegel, C. L.: Symplectic Geometry. Amer. J. Math. **65**, 1—86 (1943). — [3] Fischer, E.: Über quadratische Formen mit reellen Koeffizienten. Mh. Math. u. Phys. **16**, 234—249 (1905). — [4] Courant, R., u. D. Hilbert: Methoden der Mathematischen Physik, Bd. 1, S. 16—18. Berlin 1924.

Göttingen, Mathematisches Institut der Universität

(Eingegangen am 12. September 1955)

WIELANDT, H.
Math. Zeitschr. Bd. 63, S. 478—485 (1956)

Primitive Permutationsgruppen vom Grad $2p$

ISSAI SCHUR zum Gedächtnis

Von

HELMUT WIELANDT

1. *Übersicht.* In dieser Arbeit bedeutet p eine Primzahl und $\mathfrak{G}$ eine primitive Permutationsgruppe des Grades $2p$, welche die Symbole („Ziffern") $1, 2, \ldots, 2p$ vertauscht. $\mathfrak{G}_1$ bezeichne die größte Untergruppe $\mathfrak{G}$, welche die Ziffer 1 fest läßt. Die Anzahl der Transitivitätsgebiete, in welche die $2p$ Ziffern unter der Einwirkung von $\mathfrak{G}_1$ zerfallen, heiße t. Wir beweisen:

(1.1) *Stets ist* $t \leq 3$.

(1.2) *Ist* $t=3$, *so hat der Grad die Form* $2p=m^2+1$ $(m=1, 3, 5, \ldots)$. *Die drei transitiven Konstituenten von* $\mathfrak{G}_1$ *haben die Grade* 1, $m(m-1)/2$, $m(m+1)/2$. *Als Gruppe von komplexen linearen Transformationen betrachtet enthält* $\mathfrak{G}$ *drei irreduzible Bestandteile mit den Graden* 1, $p-1$, p.

Aus (1.1) und (1.2) folgt unmittelbar das Hauptergebnis:

Jede primitive Gruppe vom Grad $2p \neq m^2+1$ *ist zweifach transitiv.*

Offen bleibt die Frage, ob zu jedem Grade $2p=m^2+1$ primitive, aber nicht zweifach transitive Gruppen existieren. Für den kleinsten in Betracht kommenden Wert $p=5$ gibt es solche Gruppen; man erhält zwei Beispiele, wenn man die symmetrische und die alternierende Gruppe des Grades 5 auf die 10 ungeordneten Paare wirken läßt, die aus 5 Ziffern gebildet werden können.

Der Beweis von (1.1) beruht auf einem Imprimitivitätskriterium (2.2), welches vom Verfasser mit Hilfe der Methode von SCHUR [6] gefunden worden ist, aber hier auf einem anderen, elementareren Weg bewiesen wird; im übrigen werden die Methoden benutzt, mit denen BURNSIDE die zweifache Transitivität der nicht auflösbaren Gruppen vom Grade p nachgewiesen hat [1, S. 341]. Der Satz (1.2) ergibt sich durch Betrachtung der Eigenwerte einer gewissen linearen Substitution, die mit allen Elementen von $\mathfrak{G}$ vertauschbar ist.

2. *Zwei Imprimitivitätskriterien.* Im folgenden bedeutet n eine beliebige natürliche Zahl >1 und Ω die Menge der Ziffern $1, 2, \ldots, n$. Ist $\mathfrak{H}$ eine Gruppe von Permutationen dieser Ziffern und $\alpha \in \Omega$, so bezeichnen wir mit $\mathfrak{H}_\alpha$ die Untergruppe aller $H \in \mathfrak{H}$, die α ungeändert lassen. Das Bild von $\omega \in \Omega$ nach Ausführung von H bezeichnen wir mit ω^H. Entsprechend bezeichnen wir, wenn $\Gamma \subseteq \Omega$ ist, mit $\Gamma^{\mathfrak{H}}$ die Menge aller γ^H $(\gamma \in \Gamma, H \in \mathfrak{H})$; so bedeutet $\alpha^{\mathfrak{H}}$ das Transitivitätsgebiet von $\mathfrak{H}$, welchem α angehört. Die identische Per-

mutation bezeichnen wir mit E. Die Anzahl der Elemente einer beliebigen Menge M schreiben wir $|M|$; z.B. bedeutet $|\mathfrak{H}|$ die Ordnung von $\mathfrak{H}$, und $\mathfrak{H}$ ist genau dann transitiv auf Ω, wenn $|\alpha^{\mathfrak{H}}| = n$ ist.

(2.1) *Sei $\alpha \in \Omega$, $\beta \in \Omega$, $\alpha \neq \beta$ und $\Gamma \subseteq \Omega$, $\Delta < \Omega$, $|\Gamma| \geq |\Delta|$. Sei $\mathfrak{H}$ eine auf Ω transitive Permutationsgruppe zusammengesetzter Ordnung. Aus $H \in \mathfrak{H}$ und $\alpha^H \in \Gamma$ folge stets $\beta^H \in \Delta$. Dann ist $\mathfrak{H}$ imprimitiv.*

Beweis. Wir bilden die Menge $\mathfrak{C}$ aller $C \in \mathfrak{H}$, für die $\alpha^C \in \Gamma$ ist, und die Menge $\mathfrak{D}$ aller $D \in \mathfrak{H}$, für die $\beta^D \in \Delta$ ist. Nach Voraussetzung ist $\mathfrak{C} \subseteq \mathfrak{D}$. Ferner gilt $\mathfrak{H}_\alpha \mathfrak{C} = \mathfrak{C}$, $\mathfrak{H}_\beta \mathfrak{D} = \mathfrak{D}$. Hieraus folgt wegen der Transitivität von $\mathfrak{H}$

$$|\mathfrak{C}| = |\mathfrak{H}_\alpha|\,|\Gamma| \geq |\mathfrak{H}_\alpha|\,|\Delta| = |\mathfrak{H}_\beta|\,|\Delta| = |\mathfrak{D}|.$$

Wir haben nun $|\mathfrak{C}| \geq |\mathfrak{D}|$ und $\mathfrak{C} \subseteq \mathfrak{D}$, also $\mathfrak{C} = \mathfrak{D}$. Wenn wir also die Gruppe $\mathfrak{M}$ aller $M \in \mathfrak{H}$ bilden, welche die Eigenschaft $M\mathfrak{D} = \mathfrak{D}$ haben, so ist $\mathfrak{H}_\alpha \subseteq \mathfrak{M}$ und $\mathfrak{H}_\beta \subseteq \mathfrak{M}$. Wegen $\Delta \neq \Omega$ ist $\mathfrak{M} \neq \mathfrak{H}$. Nehmen wir nun entgegen der Behauptung von (2.1) an, $\mathfrak{H}$ sei primitiv. Dann sind $\mathfrak{H}_\alpha$ und $\mathfrak{H}_\beta$ maximal in $\mathfrak{H}$, daher ist $\mathfrak{H}_\alpha = \mathfrak{M} = \mathfrak{H}_\beta$. Hiernach enthält der Normalisator von $\mathfrak{H}_\alpha$ nicht nur $\mathfrak{H}_\alpha$, sondern auch alle Elemente $H \in \mathfrak{H}$ mit $\alpha^H = \beta$; er ist also größer als $\mathfrak{H}_\alpha$ und daher gleich $\mathfrak{H}$. Dann ist aber $\mathfrak{H}_\alpha = E$, also E maximal in $\mathfrak{H}$ entgegen der Voraussetzung, daß $|\mathfrak{H}|$ keine Primzahl ist. Damit ist (2.1) bewiesen. (Setzt man darin $\Gamma = \Delta$, so wird, wie man leicht erkennt, die Voraussetzung über $|\mathfrak{H}|$ entbehrlich, und (2.1) geht in einen bekannten Satz von RUDIO [4] über.)

Aus (2.1) gewinnen wir das zu Beginn angekündigte Kriterium:

(2.2) *Sei $\mathfrak{H}$ eine auf Ω transitive Permutationsgruppe zusammengesetzter Ordnung. Es gebe in Ω zwei Ziffern $\alpha \neq \beta$ und in $\mathfrak{H}$ eine Untergruppe $\mathfrak{U}$ derart, daß*

$$\beta^{\mathfrak{H}_\alpha} \subseteq \beta^{\mathfrak{U}} < \Omega \quad \text{und} \quad |\alpha^{\mathfrak{U}}| \geq |\beta^{\mathfrak{U}}|$$

gilt. Dann ist $\mathfrak{H}$ imprimitiv.

Beweis. Wir setzen $\alpha^{\mathfrak{U}} = \Gamma$, $\beta^{\mathfrak{U}} = \Delta$. Dann sind die Voraussetzungen von (2.1) erfüllt. Denn aus $\alpha^H \in \Gamma$ folgt $\alpha^H = \alpha^U$ mit einem geeigneten $U \in \mathfrak{U}$; es ist $HU^{-1} \in \mathfrak{H}_\alpha$, also

$$\beta^H = \beta^{HU^{-1}U} = (\beta^{HU^{-1}})^U \in \beta^{\mathfrak{H}_\alpha U} \subseteq \beta^{\mathfrak{U}U} = \beta^{\mathfrak{U}} = \Delta.$$

3. *Beschränkung auf gerade Permutationen.* Wir wenden uns nun der zu Beginn genannten primitiven Gruppe $\mathfrak{G}$ des Grades $n = 2p$ zu. Da die zu beweisenden Sätze (1.1) und (1.2) für $p = 2$ trivial sind, nehmen wir $p \geq 3$ an. Ferner können und wollen wir voraussetzen, daß $t \geq 3$, also $\mathfrak{G}$ nicht zweifach transitiv ist; sonst ist nichts mehr zu beweisen. Unter diesen Voraussetzungen behaupten wir:

(3.1) *Die von allen Elementen P der Ordnung p in $\mathfrak{G}$ erzeugte Untergruppe $\mathfrak{G}^*$ hat die folgenden Eigenschaften:*

(a) *$\mathfrak{G}^*$ enthält nur gerade Permutationen.*

(b) *$\mathfrak{G}^*$ ist auf Ω transitiv.*

(c) $\mathfrak{G}^*$ *ist primitiv.*

(d) *Die Anzahl der Transitivitätsgebiete von* $\mathfrak{G}_1^*$ *ist* $t^* \geq t \geq 3$.

Beweis. (a) Für jedes erzeugende Element ist $P = P^{p+1}$. Wegen $2|p+1$ ist P gerade.

(b) p teilt n und daher auch $|\mathfrak{G}|$, also ist $\mathfrak{G}^* \neq E$; ferner ist $\mathfrak{G}^*$ normal in der primitiven Gruppe $\mathfrak{G}$. Daher ist $\mathfrak{G}^*$ transitiv.

(c) Nehmen wir an, $\mathfrak{G}^*$ sei imprimitiv. Dann gibt es zwischen $\mathfrak{G}_1^*$ und $\mathfrak{G}^*$ eine eigentliche Zwischengruppe: $\mathfrak{G}_1^* < \mathfrak{R} < \mathfrak{G}^*$. Wegen der Transitivität von $\mathfrak{G}^*$ ist der Index $|\mathfrak{G}^* : \mathfrak{G}_1^*| = 2p$, also ist entweder $|\mathfrak{G}^* : \mathfrak{R}| = 2$ oder $|\mathfrak{R} : \mathfrak{G}_1^*| = 2$. Das erste kann nicht eintreten, denn sonst würde $\mathfrak{R}$ als ein Normalteiler des Index 2 in $\mathfrak{G}^*$ alle Elemente $P^{p+1} = P$ enthalten. Folglich ist $|\mathfrak{R} : \mathfrak{G}_1^*| = 2$, und der Normalisator $\mathfrak{L}$ von $\mathfrak{G}_1^*$ in $\mathfrak{G}$ enthält nicht nur, wie sich von selbst versteht, die Gruppe $\mathfrak{G}_1$, sondern $\mathfrak{L}$ enthält auch die nicht in $\mathfrak{G}_1$ liegende Gruppe $\mathfrak{R}$. Wegen der Maximalität von $\mathfrak{G}_1$ ist also $\mathfrak{L} = \mathfrak{G}$. Nun folgt aus der Transitivität von $\mathfrak{G}$, daß $\mathfrak{G}_1^* = E$, also $\mathfrak{G}^*$ regulär ist. Dann besteht aber jedes Element der Ordnung 2 in $\mathfrak{G}^*$ aus p Transpositionen, ist also ungerade, im Widerspruch zu (a).

(d) ist trivial, da $\mathfrak{G}_1^*$ Untergruppe von $\mathfrak{G}_1$ ist.

Nach (3.1) genügt es, den Beweis der Sätze (1.1) und (1.2) für $\mathfrak{G}^*$ statt $\mathfrak{G}$ zu führen. Zur Vereinfachung lassen wir den Stern fort und machen o.B.d.A. die zusätzliche Voraussetzung:

(3.2) $\mathfrak{G}$ *enthält nur gerade Permutationen und wird von Elementen der Ordnung* p *erzeugt.*

4. *Die* p*-Sylow-Normalisatoren.* Unser Beweis des Satzes (1.1) erfordert ein genaues Studium des Normalisators $\mathfrak{N}$ einer p-Sylow-Gruppe $\mathfrak{P}$ von $\mathfrak{G}$. Wir zeigen zunächst:

(4.1) *Kein Element von* $\mathfrak{G}$ *ist ein Zyklus der Ordnung* p.

Beweis. Sonst müßte $\mathfrak{G}$ als eine primitive Gruppe, die eine primitive Untergruppe kleineren Grades enthält, nach einem Satz von JORDAN [3, S. 384] zweifach transitiv sein.

(4.2) $\mathfrak{P}$ *besitzt genau zwei Transitivitätsgebiete* Γ, Δ. *Es ist* $|\Gamma| = |\Delta| = |\mathfrak{P}| = p$.

Dies folgt aus (4.1) in Verbindung mit $n = 2p$.

(4.3) *Kein Element* A *von* $\mathfrak{G}$ *hat eine Ordnung* pq *mit* $q > 1$.

Beweis. Sonst hat A^q die Ordnung p, besteht also nach (4.2) aus zwei Zyklen der Ordnung p. Daher treten in der Zyklenzerlegung von A nur solche Zyklen auf, deren Ordnungen Vielfache von p (aber nicht alle gleich p) sind. Wegen $n = 2p$ kann A nur aus einem einzigen Zyklus der Ordnung $2p$ bestehen. Dann ist aber A ungerade, im Widerspruch zu (3.2).

(4.4) *Es ist* $\Gamma^{\mathfrak{N}} \subseteq \Gamma$ *und* $\Delta^{\mathfrak{N}} \subseteq \Delta$.

Beweis. Sonst gäbe es ein $N \in \mathfrak{N}$, welches Γ mit Δ vertauscht; nach (3.2) wäre N gerade. Sei P ein erzeugendes Element von $\mathfrak{P}$ und $N^{-1}PN = P'$.

Wir wählen eine gerade Permutation R (nicht notwendig in $\mathfrak{G}$), für welche ebenfalls $R^{-1}PR=P'$, aber außerdem $\Gamma^R=\Gamma$ gilt. Dann wäre NR^{-1} eine gerade Permutation, welche mit P vertauschbar ist und außerdem Γ mit Δ vertauscht. Eine solche gibt es aber nicht.

(4.5) *Sei* $|\mathfrak{N}|=pf$. *Dann besitzt* $\mathfrak{N}_1$ *außer zwei Transitivitätsgebieten von je einer Ziffer nur noch Transitivitätsgebiete von je* f *Ziffern.*

Beweis. Nach (4.2) und (4.4) ist $|1^{\mathfrak{N}}|=p$, daher ist $|\mathfrak{N}_1|=f$. Sei $1\in\Gamma$. Kein Element von $\mathfrak{N}$ hat auf Γ den identischen Konstituenten (E ausgenommen); denn es müßte $\mathfrak{P}$ zentralisieren, also aus einem Zyklus der Ordnung p auf Δ bestehen, entgegen (4.1). Hiernach hat der Konstituent von $\mathfrak{N}_1$ auf Γ die Ordnung f und besitzt daher, weil er im Normalisator der von einem p-Zyklus erzeugten Gruppe liegt, außer einem Transitivitätsgebiet der Länge 1 nur solche der Länge f. Das gleiche gilt für den Konstituenten von $\mathfrak{N}_1$ auf Δ.

5. *Abschätzung von* t. Als Vorstufe zu (1.1) beweisen wir:

(5.1) *Es ist* $t<2+(p-1)/f$.

Beweis. (a) Ist $\beta\in\Omega$, $\beta\neq 1$, so ist $|\beta^{\mathfrak{G}_1}|>1$. Das ergibt sich z. B. dadurch, daß man in (2.2) $\mathfrak{H}=\mathfrak{G}$, $\alpha=1$ und $\mathfrak{U}=\mathfrak{G}_1$ setzt.

(b) Nach (a) besteht jedes von 1 verschiedene Transitivitätsgebiet von $\mathfrak{G}_1$ aus mindestens zwei Ziffern. Daraus folgt $t<p+1$, womit insbesondere der Fall $f=1$ erledigt ist. Ist $f>1$, so ziehen wir Satz (4.5) heran. Er besagt, daß $\mathfrak{N}_1$ außer 1 noch eine weitere Ziffer fest läßt (etwa β) und sonst nur Transitivitätsgebiete aus je f Ziffern besitzt. Nach (4.4) ist jedes Transitivitätsgebiet von $\mathfrak{N}_1$ ganz in einem Transitivitätsgebiet von $\mathfrak{P}$ enthalten. Die Transitivitätsgebiete von $\mathfrak{G}_1$ seien $T_1=1$, $T_2\ni\beta$, $T_3,\ldots,T_t$. Dann gilt:

(c) Es ist $|T_2|>1$ nach (a).

(d) Es ist $|T_2|\geq f+1$. Denn es gibt $\gamma\in T_2$, $\gamma\neq\beta$, und damit haben wir $\gamma^{\mathfrak{N}_1}\subseteq\gamma^{\mathfrak{G}_1}=T_2$. Nach (4.5) ist $|\gamma^{\mathfrak{N}_1}|=f$.

(e) Für $\tau\geq 3$ ist $|T_\tau|\geq 2f$. Denn wegen $\mathfrak{N}_1\subseteq\mathfrak{G}_1$ besteht T_τ aus vollen Transitivitätsgebieten von $\mathfrak{N}_1$; von diesen hat jedes die Länge f; T_τ besteht nicht nur aus einem einzigen Transitivitätsgebiet von $\mathfrak{N}_1$, sonst wäre $\mathfrak{G}$ wieder imprimitiv [in (2.2) wähle man $\alpha=1$, $\beta\in T_\tau$, $\mathfrak{H}=\mathfrak{G}$, $\mathfrak{U}=\mathfrak{P}$].

(f) Aus den Aussagen (d) und (e) folgt

$$1+(f+1)+(t-2)\cdot 2f\leq\sum_{\tau=1}^{t}|T_\tau|=2p.$$

Hieraus ergibt sich die Behauptung von Satz (5.1): Es ist

$$t-2\leq(2p-2-f)/2f<(2p-2)/2f=(p-1)/f.$$

6. *Die irreduziblen Bestandteile von* $\mathfrak{G}$. Wir sehen die Elemente $G\in\mathfrak{G}$ nun nicht mehr als Permutationen auf Ω an, sondern als $2p$-reihige Matrizen, die in jeder Zeile und in jeder Spalte genau einmal 1 und sonst überall 0 haben. Die absolut irreduziblen Bestandteile von $\mathfrak{G}$ nennen wir $\mathfrak{D}_1,\mathfrak{D}_2,\ldots,\mathfrak{D}_r$, ihre Grade $f_1,f_2,\ldots,f_r$ und die Spur der Matrix $\mathfrak{D}_\varrho(G)$ den Charakter $\chi_\varrho(G)$.

Die Numerierung sei so gewählt, daß $\mathfrak{D}_1$ die identische Darstellung ist (daher $f_1=1$). Da $\mathfrak{G}$ transitiv ist, ist für $\varrho>1$ stets $\mathfrak{D}_\varrho \neq \mathfrak{D}_1$, und da $\mathfrak{G}$ nicht zweifach transitiv ist, ist $r>2$ [1, S. 275]. Mit ζ bezeichnen wir eine fest gewählte primitive p-te Einheitswurzel, mit $\mathfrak{P}$ eine bestimmte p-Sylow-Gruppe von $\mathfrak{G}$.

(6.1) *Ist* $E \neq P \in \mathfrak{P}$, *so hat* P, *als Matrix betrachtet, die Eigenwerte* $1, \zeta, \ldots, \zeta^{p-1}$, *jeden genau zweifach.*

Beweis. P zerfällt nach (4.2) in zwei Zyklen der Ordnung p.

Nach (6.1) gibt es genau ein $\varrho>1$, für welches die Matrix $\mathfrak{D}_\varrho(P)$ den Eigenwert 1 besitzt. Die Numerierung sei so gewählt, daß $\varrho=2$ ist. Wir behaupten:

(6.2) $\chi_2(G)$ *ist rational für jedes* $G \in \mathfrak{G}$.

Beweis. Da die Spuren von $\mathfrak{G}$ rational sind, enthält $\mathfrak{G}$ neben jedem Charakter auch alle konjugierten Charaktere. Beim Übergang zu den algebraisch konjugierten Darstellungen muß $\mathfrak{D}_2$ bis auf Äquivalenz fest bleiben, da $\mathfrak{D}_2$ die einzige in $\mathfrak{G}$ enthaltene irreduzible Darstellung $\neq \mathfrak{D}_1$ ist, in welcher P den rationalen Eigenwert 1 hat.

(6.3) *Es ist* $f_2=p$. *Die Eigenwerte von* $\mathfrak{D}_2(P)$ *sind* $1, \zeta, \ldots, \zeta^{p-1}$.

Beweis. Sei $k \geq 0$ die Vielfachheit, mit der ζ als Eigenwert von $\mathfrak{D}_2(P)$ auftritt. Da $\mathfrak{D}_2$ zu den algebraisch konjugierten Darstellungen äquivalent ist, besitzt $\mathfrak{D}_2(P)$ jeden Wert $\zeta^\nu \neq 1$ als Eigenwert der Vielfachheit k. Es ist also $f_2=1+k(p-1)$. Wegen $r>2$ haben wir $f_2<2p-1$, also $k<2$. Andererseits ist $k>0$; sonst würden $\mathfrak{D}_2$ und $\mathfrak{D}_1$ auf $\mathfrak{P}$ übereinstimmen, also auch auf den zu $\mathfrak{P}$ konjugierten Untergruppen; da diese nach (3.2) die ganze Gruppe $\mathfrak{G}$ erzeugen, wäre $\mathfrak{D}_2=\mathfrak{D}_1$. Es bleibt nur die Möglichkeit $k=1$, also $f_2=p$.

(6.4) *Sei* $3 \leq \varrho < \sigma \leq r$. *Dann sind* $\mathfrak{D}_\varrho$ *und* $\mathfrak{D}_\sigma$ *nicht äquivalent; aber für jedes Element* A *in* $\mathfrak{G}$, *dessen Ordnung* $a \neq p$ *ist, gilt* $\chi_\varrho(A)=\chi_\sigma(A)$. *Insbesondere ist* $f_\varrho=f_\sigma$.

Beweis. Wir setzen $\chi_2+\chi_4+\cdots+\chi_r=\chi$ und bezeichnen eine zugehörige Darstellung von $\mathfrak{G}$ mit $\mathfrak{D}$. Nach (6.1) sind alle Eigenwerte von $\mathfrak{D}(P)$ voneinander verschieden. Daher sind $\mathfrak{D}_\varrho$ und $\mathfrak{D}_\sigma$ nicht äquivalent. Sei ζ^k ein Eigenwert von $\mathfrak{D}_\varrho(P)$ und ζ^l einer von $\mathfrak{D}_\sigma(P)$. Nach (4.3) ist $(p, a)=1$, also besitzt der Körper der $|\mathfrak{G}|$-ten Einheitswurzeln einen Automorphismus ψ mit $\psi[\zeta^k]=\zeta^l$ und $\psi[\chi_\varrho(A)]=\chi_\varrho(A)$. Der Charakter χ enthält den Charakter $\psi[\chi_\varrho]$, da er rational ist und χ_ϱ enthält. In der zu $\psi[\chi_\varrho]$ gehörigen Darstellung besitzt P den Eigenwert ζ^l. Da $\mathfrak{D}_\sigma$ der einzige Bestandteil von $\mathfrak{D}$ ist, in dem P diesen Eigenwert hat, ist $\psi[\chi_\varrho]=\chi_\sigma$. Insbesondere ist $\chi_\varrho(A)=\psi[\chi_\varrho(A)]=\chi_\sigma(A)$, wie behauptet.

(6.5) *Es ist* $t=r$.

Beweis. In jeder transitiven Permutationsgruppe $\mathfrak{H}$ ist die Anzahl der Transitivitätsgebiete von $\mathfrak{H}_1$ gleich der Quadratsumme der Vielfachheiten der

irreduziblen Bestandteile von $\mathfrak{H}$ [1, S. 275]. Da die r irreduziblen Bestandteile von $\mathfrak{G}$ alle voneinander verschieden sind, hat die Quadratsumme in unserem Fall den Wert r.

7. *Beweis des Satzes* (1.1). Nehmen wir an, entgegen der Behauptung von (1.1) sei $t>3$. Dann ist nach (6.5) auch $r>3$, und wir können (6.4) auf $\varrho=3$, $\sigma=4$ anwenden. Die Orthogonalitätsrelationen ergeben

$$g=|\mathfrak{G}|=\sum_{G\in\mathfrak{G}}\chi_3(G)\{\bar{\chi}_3(G)-\bar{\chi}_4(G)\}=\sum_{\mathfrak{Q}}\sum_{Q\in\mathfrak{Q}}\chi_3(Q)\{\bar{\chi}_3(Q)-\bar{\chi}_4(Q)\}.$$

Hierbei durchläuft $\mathfrak{Q}$ alle p-Sylow-Gruppen von $\mathfrak{G}$. Das sind die g/pf zu $\mathfrak{P}$ konjugierten Gruppen; daher liefert jede der inneren Summen den gleichen Beitrag und wir erhalten

$$g=\frac{g}{pf}\sum_{P\in\mathfrak{P}}\chi_3(P)\{\bar{\chi}_3(P)-\bar{\chi}_4(P)\}=\frac{g}{pf}\sum_{P\in\mathfrak{P}}|\chi_3(P)|^2=\frac{g}{pf}\cdot pf_3.$$

Hieraus folgt $f_3=f$, also nach (6.4) weiter $f_\varrho=f$ $(3\leq\varrho\leq r)$. Nun ergibt sich für den Grad nach (6.3) und (6.5) die Zerlegung

$$2p=1+p+(r-2)f=1+p+(t-2)f.$$

Hieraus folgt $t=2+(p-1)/f$, im Widerspruch zu (5.1). Damit ist (1.1) bewiesen.

8. *Beweis des Satzes* (1.2). Von nun an sei $t=3$, also auch $r=3$.

(8.1) *Die beiden von* 1 *verschiedenen Transitivitätsgebiete von* $\mathfrak{G}_1$ *seien* Ξ, H. *Es sei* $|\Xi|=x$, $|H|=y$ *und etwa* $x\leq y$. *Dann ist* (a) $2p=1+x+y$ *und* (b) $2<x<p$.

Beweis. Die Behauptung (a) folgt aus $\Omega=\{1\}\cup\Xi\cup H$.

(b) Nach dem Beweisteil (a) von Satz (5.1) ist $x>1$. Nehmen wir $x=2$ an, also etwa $\Xi=\{\alpha,\beta\}$. Dann hat die Gruppe $\mathfrak{G}_{1\alpha}$ in $\mathfrak{G}_1$ den Index 2, ebenso in $\mathfrak{G}_\alpha$; sie ist also normal in beiden Gruppen und daher auch in deren Erzeugnis $\mathfrak{G}$; hieraus folgt $\mathfrak{G}_{1\alpha}=E$, also $|\mathfrak{G}_1|=2$, $y=2$, $2p=5$. Dieser Widerspruch zeigt, daß $x>2$ ist. Andererseits ist $x<p$; sonst wäre $y\geq x\geq p$ und nach (a) $2p\geq 1+p+p$.

(8.2) *Es gibt eine quadratische Matrix* X, *welche in jeder ihrer* $2p$ *Spalten an genau* x *Stellen mit* 1, *sonst überall mit* 0 *besetzt ist und mit jeder Matrix aus* $\mathfrak{G}$ *vertauschbar ist. Es ist* $\mathrm{Spur}\,X=0$ *und* $\mathrm{Spur}\,\bar{X}'X=2px$.

Beweis. Man konstruiert X nach Burnside [2, S. 174] und Schur [6, S. 599], indem man in der ersten Spalte diejenigen Stellen mit 1 besetzt, deren Zeilenindizes in Ξ liegen. Die übrigen Spalten von X sind dann durch die Forderung der Vertauschbarkeit mit allen $G\in\mathfrak{G}$ festgelegt. Dabei erhalten alle Diagonalelemente von X den gleichen Wert, und zwar den Wert 0 wegen $1\notin\Xi$. Daher ist $\mathrm{Spur}\,X=0$. Die Spur von $\bar{X}'X$ ist die Absolutquadratsumme aller Elemente von X, also gleich $2px$.

(8.3) *Aus der in* (8.2) *genannten Matrix X werde mittels einer unitären Matrix U, für welche $U\mathfrak{G}U^{-1}$ vollständig reduziert ist, die Matrix $Z = UXU^{-1}$ gebildet. Dann ist Z eine Diagonalmatrix, in welcher ein Diagonalelement z_1 einfach, ein weiteres Diagonalelement z_2 p-fach und ein drittes Diagonalelement z_3 $(p-1)$-fach auftritt.*

Beweis. Die Gruppe $U\mathfrak{G}U^{-1}$ ist vollständig reduziert, ihre irreduziblen Bestandteile $\mathfrak{D}_\varrho$ $(\varrho = 1, 2, 3)$ sind paarweise nicht äquivalent und haben die Grade $1, p, p-1$. Jede Matrix dieser Gruppe ist mit Z vertauschbar. Nach dem bekannten Lemma von SCHUR [5, S. 409] hat Z die angegebene Gestalt.

(8.4) *Die Zahlen z_1, z_2, z_3 sind ganz rational. Es gilt*

(a) $z_1 + p z_2 + (p-1) z_3 = 0$

(b) $|z_1|^2 + p|z_2|^2 + (p-1)|z_3|^2 = 2px$

(c) $z_2 \neq x, \quad z_3 \neq x$

(d) $z_1 = x$.

Beweis. Die linke Seite von (a) hat den Wert $\mathrm{Spur}\, Z = \mathrm{Spur}\, X = 0$. Die linke Seite von (b) ist $\mathrm{Spur}\, \bar{Z}'Z = \mathrm{Spur}\, \bar{X}'X = 2px$. Wäre $z_2 = x$, so würde aus (b) folgen, daß $px^2 \leq 2px$ ist, also $x \leq 2$, entgegen (8.1). Wäre $z_3 = x$, so würden wegen (a) jedenfalls z_1 und z_2 nicht beide verschwinden, also ergäbe sich aus (b) die mit (8.1) unverträgliche Ungleichung $(p-1)x^2 < 2px$. Mindestens eine der Zahlen z_i muß aber mit x übereinstimmen. Denn x ist ein Eigenwert von X, da für die Zeile $\mathfrak{e} = (1, 1, \ldots, 1)$ nach Konstruktion von X die Gleichung $\mathfrak{e}X = x\mathfrak{e}$ gilt. Nach (c) ist $z_1 = x$. Insbesondere ist z_1 ganz rational.

Daß auch z_2 und z_3 ganz rational sind, ergibt sich so: Die charakteristische Gleichung von X hat ganze rationale Koeffizienten. Daher sind die Wurzeln z_2 und z_3 zunächst einmal ganz algebraisch. Ferner ist die genaue Vielfachheit der Wurzel z_3 entweder $p-1$ oder, wenn $z_2 = z_3$ sein sollte, $2p-1$; in jedem Fall ist sie verschieden von der Vielfachheit jeder anderen Wurzel der charakteristischen Gleichung. Folglich ist z_3 rational. Daß auch z_2 rational ist, folgt nun aus (a).

Wir kommen zum Beweis des Satzes (1.2). Lesen wir die unter (8.4a) bewiesene Gleichung $\mathrm{mod}\, p$, so erhalten wir wegen $z_1 = x$ die Kongruenz

(e) $$x - \bar{z}_3 \equiv 0 \bmod p.$$

Ferner ergibt sich aus (b) die Abschätzung

$$z_3^2 \leq 2px/(p-1) \leq px < p^2.$$

Hiernach ist $|z_3| < p$ und wegen $0 < x < p$ weiter

(f) $$-p < x - z_3 < 2p.$$

Durch (c), (e), (f) ist $x - z_3 = p$ eindeutig bestimmt. Nun können wir vermöge (a) und (d) sowohl x wie z_3 durch z_2 ausdrücken. Das Ergebnis ist

(g) $$x = p - z_2 - 1, \quad z_3 = -z_2 - 1.$$

Führen wir diese Werte in (b) ein, so ergibt sich

$$z_2^2 + (z_2 + 1)^2 = p.$$

Setzen wir $2z_2 + 1 = m$, so folgt aus der letzten Gleichung

$$2p = m^2 + 1.$$

Aus (g) folgt nun

$$x = \frac{m(m-1)}{2}$$

und hieraus wegen $1 + x + y = 2p$ weiter

$$y = \frac{m(m+1)}{2}.$$

Schließlich erkennen wir aus den beiden letzten Gleichungen, daß $m = y - x \geq 0$ ist, und nach Definition ist $m = 2z_2 + 1$ ungerade; m ist also eine Zahl aus der Reihe $1, 3, 5, \ldots$. Damit sind alle Behauptungen des Satzes (1.2) bewiesen.

Literatur

[1] Burnside, W.: Theory of groups of finite order, 2. Aufl. Cambridge University Press 1911. Dover Publications, New York 1955. — [2] Burnside, W.: On some properties of groups of odd order. Proc. Lond. Math. Soc. **33**, 162—185 (1901). — [3] Jordan, C.: Théorèmes sur les groupes primitifs. J. Math. pures appl. **16**, 383—408 (1871). — [4] Rudio, F.: Über primitive Gruppen. J. reine u. angew. Math. **102**, 1—8 (1888). — [5] Schur, I.: Neue Begründung der Theorie der Gruppencharaktere. Sitzgsber. preuß. Akad. Wiss. **1905**, 406—432. — [6] Schur, I.: Zur Theorie der einfach transitiven Permutationsgruppen. Sitzgsber. preuß. Akad. Wiss., phys.-math. Kl. **1933**, 598—623.

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 6. September 1955)

RADO, R.
Math. Zeitschr. Bd. 63, S. 486—495 (1956)

Minimal Points of Convex Sets in Sequence Spaces

In memoriam ISSAI SCHUR

By

RICHARD RADO

1. By means of applications of the Hahn-Banach theorem W. W. ROGOSINSKI has obtained new results in the theory of linear equations in infinitely many variables and, in particular, the following theorem ([*1*], Theorem B).

THEOREM 1. *Let m be a positive integer and, for $0 \leq r < m$; $\lambda \geq 0$, $a_{r\lambda}$ and a_r be complex numbers such that the sequences $(a_{r0}, a_{r1}, \ldots)$ are linearly independent and*

$$\sum_{\lambda=0}^{\infty} |a_{r\lambda}| < \infty \qquad (r < m). \tag{1}$$

Let X be the set of all bounded sequences $x = (x_0, x_1, \ldots)$ of complex numbers such that

$$\sum_{\lambda} a_{r\lambda} x_\lambda + a_r = 0 \qquad (r < m). \tag{2}$$

Then there exists an index λ_0 and a number ξ such that, for every x of X,

$$\sup_{\lambda} |x_\lambda| \geq |\xi|, \tag{3}$$

with equality for at least one x of X; also, whenever there is equality in (3) for some x of X, then $x_{\lambda_0} = \xi$.

After being told of this result the present author noticed that it could also be deduced by means of the methods of classical analysis. This note is concerned with more general results obtained by developing such methods further. Rogosinski's methods yield generalisations of Theorem 1 in directions different from those considered here.

2. Before formulating the first generalization of Theorem 1 I introduce some definitions. The set of all sequences of complex numbers is denoted by S. The members of the sequence x will always be denoted by $x_0, x_1, \ldots$, and similarly in the case of y, z, etc. If N is a non-empty set of non-negative integers then I put[1])

$$\|x\|_N = \sup\,[\lambda \in N]\, |x_\lambda|,$$

and if N is empty, i.e. $N = 0$, then I put $\|x\|_N = 0$. Also, $\|x\|_N = \|x\|$, if $N = \{0, 1, \ldots\}$. The statement

$$x = y \quad \text{in } M$$

means that $x_\lambda = y_\lambda$ for $\lambda \in M$.

[1]) I write $\sup\,[\lambda \in N]\ |x_\lambda|$ instead of $\sup_{\lambda \in N} |x_\lambda|$, and similarly in other cases.

THEOREM 2. *Let m be a positive integer and let a_ν be complex numbers and A a $(m\times\infty)$-matrix whose elements $a_{\nu\lambda}$ $(\nu<m;\ \lambda=0,1,\ldots)$ are complex numbers satisfying* (1). *Let X be the set of all sequences x of S such that* (2) *holds and the series in* (2) *are absolutely convergent. Suppose that $X\neq 0$. Then there is a sequence z of X, a positive integer n and a partition of the set $\{0,1,\ldots\}$ into disjoint non-empty sets $N_0, N_1, \ldots, N_{n-1}$ such that the following conditions hold.*

$$|z_\lambda| = \varrho_\nu \qquad (\nu<n;\ \lambda\in N_\nu);$$

$$\varrho_0 > \varrho_1 > \cdots > \varrho_{n-1}.$$

If $x\in X-\{z\}$, then there is $\nu<n$ such that $x=z$ in $N_0+\cdots+N_{\nu-1}$,

$$\|x\|_{N_\nu+\cdots+N_{n-1}} > \varrho_\nu.$$

Furthermore, $1\leq n\leq 1+RkA$, where RkA denotes the rank of the matrix A. There is exactly one sequence z of X corresponding to which, with any value of n, sets N_ν and numbers ϱ_ν can be found having the properties stated.

It can be seen that if $RkA=m$ then, for any choice of numbers a_ν, $X\neq 0$.

In order to deduce Theorem 1 from Theorem 2 choose any element λ_0 of N_0 and put $\xi=z_{\lambda_0}$. Then the assertion of Theorem 1 holds. For, let $x\in X-\{z\}$ and let ν have the property stated in Theorem 2. If $\nu=0$, then $\|x\|>\varrho_0=|\xi|=\|z\|$. If $\nu>0$ then $x=z$ in N_0; $x_{\lambda_0}=z_{\lambda_0}=\xi$. This proves Theorem 1.

3. I introduce some further definitions. ω_0 and ω_1 denote the first denumerable and the first non-denumerable ordinal number respectively. For any ordinals k, l I denote by $[k,l]$ the set of all ordinals λ in the range $k\leq\lambda<l$. Thus

$$[k,l] = \{\lambda : k\leq\lambda<l\}.$$

Corresponding to every sequence x I define a, possibly transfinite, sequence of successive maxima ϱ_ν of x and some related concepts.

Consider a fixed x of S. Denote by R the set of all non-negative integers λ_0 which have the property that, for suitable non-negative integers $\lambda_1, \lambda_2, \ldots$,

$$|x_{\lambda_0}| < |x_{\lambda_1}| < \cdots,$$

and by L the set of all non-negative integers outside R. In symbols:

$$R = \{\lambda_0 : |x_{\lambda_0}| < |x_{\lambda_1}| < \cdots\}; \qquad L = [0,\omega_0] - R.$$

Then the set $\{-|x_\lambda| : \lambda\in L\}$ is a well-ordered initial segment of the set $Q=\{-|x_\lambda| : \lambda<\omega_0\}$, ordered by magnitude, and the corresponding terminal segment $\{-|x_\lambda| : \lambda\in R\}$ of Q has no first element. Thus L is the maximal well-ordered initial segment of the ordered set Q. Now I can write

$$\{|x_\lambda| : \lambda\in L\} = \{\varrho_\nu : \nu<n\},$$

where n is some ordinal number, finite or infinite, such that

$$n<\omega_1; \qquad \varrho_\mu > \varrho_\nu \quad (\mu<\nu<n).$$

I put

$$N_\nu = \{\lambda : |x_\lambda| = \varrho_\nu\}; \qquad L_\nu = \sum [\mu < \nu]\, N_\mu;$$

$$R_\nu = [0, \omega_0] - L_\nu \qquad (\nu < n); \qquad \varrho = \|x\|_R.$$

Then $N_\nu \neq 0$; $|x_\lambda| = \|x\|_{R_\nu} = \varrho_\nu$ $(\nu < n;\ \lambda \in N_\nu)$;

$$|x_\lambda| < \varrho \qquad (\lambda \in R).$$

To complete the definitions I put

$$N_\nu = 0; \quad R_\nu = R; \quad L_\nu = L; \quad \varrho_\nu = \varrho \qquad (n \leq \nu < \omega_1).$$

Then

$$\varrho_\mu \geq \varrho_\nu \qquad (\mu < \nu < \omega_1).$$

If I want to express the fact that the numbers n, ϱ_ν, ϱ and the sets N_ν, L_ν, R_ν, L, R refer to the sequence x, I write $n(x)$, $\varrho_\nu(x)$, ...

For sets M, N the relation $M \subset N$ denotes inclusion, in the wide sense, and the symbol MN the intersection of M and N.

4. Let $z \in X \subset S$. Then z is called a *minimal sequence of* X if the following condition holds. Whenever $x \in X$ then either

(i) $x = z$ in $L(z)$, and $\|x\|_{R(z)} \geq \varrho(z)$,

or

(ii) there is $\nu < n(z)$ such that $x = z$ in $L_\nu(z)$ and

$$\|x\|_{R_\nu(z)} > \varrho_\nu(z).$$

The set X is called *weakly closed* if, whenever[2])

$$x^s \in X \,(s < \omega_0),$$

$$\lim [s \to \omega_0]\, x_\lambda^s = y_\lambda \qquad (\lambda < \omega_0),$$

then

$$y \in X.$$

The set X is called *convex* if $x', x'' \in X$ implies $x''' \in X$, where

$$x_\lambda''' = \tfrac{1}{2}(x_\lambda' + x_\lambda'') \qquad (\lambda < \omega_0).$$

It follows that, if X is weakly closed and convex, and

$$y^\tau \in X; \quad \|y^\tau\| \leq c < \infty \qquad (\tau < \omega_0),$$

then

$$y^* \in X,$$

where

$$y_\lambda^* = \sum [\tau < \omega_0]\, 2^{-\tau-1} y_\lambda^\tau \qquad (\lambda < \omega_0).$$

---

[2]) Throughout this note, upper indices do not denote exponentiation.

5. THEOREM 3. *Let $X \subset S$. Suppose that, for some real number c, the set*

$$Y = \{x : x \in X; \quad \|x\| \leq c\}$$

is non-empty, convex and weakly closed. Then

(i) *X possesses at least one minimal sequence*

(ii) *if z and w are minimal sequences of X, then*

$$\varrho_\nu(z) = \varrho_\nu(w); \quad N_\nu(z) = N_\nu(w)\ (\nu < \omega_1); \quad n(z) = n(w);$$
$$L(z) = L(w); \quad \varrho(z) = \varrho(w); \qquad z = w \text{ in } L(z).$$

It will be seen that in order to deduce Theorem 2 from Theorem 3, it suffices to show that the set X of Theorem 2 satisfies the hypothesis of Theorem 3 and that, if z is a minimal sequence of X, then $R(z) = 0$; $n(z) \leq 1 + RkA$. The proofs of these statements will be given in § 8.

6. We begin by proving (ii) of Theorem 3. Let $\nu < \omega_1$, and suppose that, for $\mu < \nu$,

$$N_\mu(z) = N_\mu(w),$$
$$z = w \quad \text{in} \quad N_\mu(z).$$

Then $R_\nu(z) = R_\nu(w)$. Since z is minimal it follows that

$$\|w\|_{R_\nu(z)} \geq \|z\|_{R_\nu(z)} = \|z\|_{R_\nu(w)}.$$

By symmetry, we obtain $\|z\|_{R_\nu(w)} \geq \|w\|_{R_\nu(z)}$. Hence $\|w\|_{R_\nu(z)} = \|z\|_{R_\nu(z)}$. Hence, since z is minimal, $w = z$ in $N_\nu(z)$. Now, the definition of $N_\nu(w)$ shows that $N_\nu(w) > N_\nu(z)$. On grounds of symmetry, $N_\nu(z) > N_\nu(w)$ follows and so $N_\nu(z) = N_\nu(w)$. We have thus proved, by induction, that, for $\nu < \omega_1$, $N_\nu(z) = N_\nu(w)$, and $z = w$ in $N_\nu(z)$. Hence $n(z) = n(w)$; $L(z) = L(w)$; $R(z) = R(w)$,

$$z = w \quad \text{in} \quad L(z).$$

Since z is minimal, this implies $\|w\|_{R(z)} \geq \|z\|_{R(z)} = \|z\|_{R(w)}$ and, by symmetry, $\|z\|_{R(w)} \geq \|w\|_{R(z)}$. Hence $\varrho(z) = \|z\|_{R(z)} = \|w\|_{R(w)} = \varrho(w)$, and (ii) is established.

7. I now proceed to prove (i) of Theorem 3. As a first step, I define values of symbols

$$x^\nu, \ N_\nu, \ L_\nu, \ R_\nu, \ \varrho_\nu \qquad (\nu < \omega_1). \tag{4}$$

I shall ultimately show that the numbers and sets (4) belong to some sequence z, in the sense defined in § 3. Let $\nu < \omega_1$. Suppose that

$$x^\mu \in Y; \quad N_\mu \subset [0, \omega_0] \qquad (\mu < \nu);$$
$$x^{\mu_0} = x^{\mu_1} \quad \text{in} \quad N_{\mu_0} \qquad (\mu_0 < \mu_1 < \nu).$$

Put

$$L_\nu = \sum [\mu < \nu]\, N_\mu; \quad R_\nu = [0, \omega_0] - L_\nu;$$
$$Y_\nu = \prod [\mu < \nu]\, \{x : x \in Y; \quad x = x^\mu \quad \text{in} \quad N_\mu\}.$$

In particular, $L_0=0$; $Y_0=Y$. I now show that $Y_\nu \neq 0$. If $\nu=0$, then $Y_\nu=Y\neq 0$. If $\nu=\tau+1$, then $x^\tau\in Y_\nu$. Now let ν be a limit ordinal. Then there is $y^\nu\in S$ and a subset

$$\{\mu_{\nu 0},\mu_{\nu 1},\dots\} \tag{5}$$

of $[0,\nu]$ such that $\mu_{\nu 0}<\mu_{\nu 1}<\cdots$,

$$\sup\,[\tau<\omega_0]\,\mu_{\nu\tau}=\nu,$$

$$\lim\,[\tau\to\omega_0]\,x_\lambda^{\mu_{\nu\tau}}=y_\lambda^\nu \qquad (\lambda<\omega_0). \tag{6}$$

Then $y^\nu\in Y$, since Y is weakly closed. Let $\mu<\nu$; $\lambda\in N_\mu$. Then there is $\sigma<\omega_0$ such that $\mu_{\nu\sigma}\geq\mu$. Then, for $\sigma\leq\tau<\omega_0$,

$$x_\lambda^\mu=x_\lambda^{\mu_{\nu\tau}};\qquad x_\lambda^\mu=y_\lambda^\nu.$$

Hence

$$y^\nu=x^\mu \text{ in } N_\mu;\qquad y^\nu\in Y_\nu.$$

Put

$$\varrho_\nu=\inf\,[x\in Y_\nu]\,\|x\|_{R_\nu};$$

$$Z_\nu=\{x: x\in Y_\nu;\ \|x\|_{R_\nu}=\varrho_\nu\}.$$

Then there is $y^{\nu\tau}\in Y_\nu$; $z^\nu\in S$ such that

$$\|y^{\nu\tau}\|_{R_\nu}<\varrho_\nu+2^{-\tau} \qquad (\tau<\omega_0),$$

$$\lim\,[\tau\to\omega_0]\,y_\lambda^{\nu\tau}=z_\lambda^\nu \qquad (\lambda<\omega_0).$$

Then $z^\nu\in Y$. If $\mu<\nu$; $\lambda\in N_\mu$, then $y_\lambda^{\nu\tau}=x_\lambda^\mu$; $z_\lambda^\nu=x_\lambda^\mu$. Hence $z^\nu=x^\mu$ in N_μ; $z^\nu\in Y_\nu$.

It follows from the definition of ϱ_ν that $\|z^\nu\|_{R_\nu}\geq\varrho_\nu$. On the other hand, if $\lambda\in R_\nu$, then

$$|y_\lambda^{\nu\tau}|\leq\|y^{\nu\tau}\|_{R_\nu}<\varrho_\nu+2^{-\tau},$$

$$|z_\lambda^\nu|\leq\varrho_\nu;\qquad \|z^\nu\|_{R_\nu}\leq\varrho_\nu.$$

Hence

$$\|z^\nu\|_{R_\nu}=\varrho_\nu;\qquad z^\nu\in Z_\nu;\qquad Z_\nu\neq 0.$$

Put

$$N_\nu=\bigsqcap\,[x\in Z_\nu]\,\{\lambda:\lambda\in R_\nu;\ |x_\lambda|=\varrho_\nu\}.$$

Now I define x^ν. For every $\tau\in R_\nu-N_\nu$ there is $x^{\nu\tau}\in Z_\nu$ such that $|x_\tau^{\nu\tau}|<\varrho_\nu$. For $\tau\in[0,\omega_0]-(R_\nu-N_\nu)$ I put $x^{\nu\tau}=z^\nu$. Then I put

$$x_\lambda^\nu=\sum\,[\tau<\omega_0]\,2^{-\tau-1}x_\lambda^{\nu\tau} \qquad (\lambda<\omega_0).$$

This completes the inductive definition of the entities (4).

If $\lambda<\omega_0$, then

$$|x_\lambda^\nu|\leq\sum\,[\tau<\omega_0]\,2^{-\tau-1}|x_\lambda^{\nu\tau}|\leq\sum\,[\tau<\omega_0]\,2^{-\tau-1}c=c.$$

Hence, since Y is convex and weakly closed, $x^\nu\in Y$. If $\mu<\nu$; $\lambda\in N_\mu$, then $x^{\nu\tau}\in Z_\nu\subset Y_\nu$; $x_\lambda^{\nu\tau}=x_\lambda^\mu$,

$$x_\lambda^\nu=\sum\,[\tau<\omega_0]\,2^{-\tau-1}x_\lambda^{\nu\tau}=x_\lambda^\mu,\qquad x^\nu=x^\mu \text{ in } N_\mu;\qquad x^\nu\in Y_\nu.$$

Hence, by definition of ϱ_ν, $||x^\nu||_{R_\nu} \geq \varrho_\nu$. If $\lambda \in R_\nu$, then $|x_\lambda^{\nu\tau}| \leq ||x^{\nu\tau}||_{R_\nu} \leq \varrho_\nu$; $|x_\lambda^\nu| \leq \varrho_\nu$,

$$||x^\nu||_{R_\nu} \leq \varrho_\nu.$$

Hence $||x^\nu||_{R_\nu} = \varrho_\nu$; $x^\nu \in Z_\nu$. If $\tau \in R_\nu - N_\nu$ then, by definition of $x^{\nu\tau}$, $|x_\tau^{\nu\tau}| < \varrho_\nu$. Hence

$$|x_\tau^\nu| \leq 2^{-\tau-1}|x_\tau^{\nu\tau}| + \sum [\sigma \neq \tau]\, 2^{-\sigma-1}|x_\tau^{\nu\sigma}| < 2^{-\tau-1}\varrho_\nu + \sum [\sigma \neq \tau]\, 2^{-\sigma-1}\varrho_\nu = \varrho_\nu,$$

$$|x_\tau^\nu| < \varrho_\nu \qquad (\tau \in R_\nu - N_\nu). \tag{7}$$

Now I show, by induction with respect to ν, that

$$\varrho_\mu \geq \varrho_\nu \qquad (\mu < \nu < \omega_1). \tag{8}$$

Assume that, for some fixed $\nu < \omega_1$ and every μ, ν_0 such that $\mu < \nu_0 < \nu$, the relation $\varrho_\mu \geq \varrho_{\nu_0}$ holds. Let $\mu < \nu$. If $\nu = \tau + 1$, then $x^\tau \in Y_\nu$,

$$\varrho_\nu \leq ||x^\tau||_{R_\nu} \leq ||x^\tau||_{R_\tau} = \varrho_\tau \leq \varrho_\mu.$$

Now let ν be a limit number. Then there is $\tau_0 < \omega_0$ such that $\mu_{\nu\tau_0} \geq \mu$, where $\mu_{\nu\tau}$ is the number occurring in (5). Then, for $\lambda \in R_\nu$; $\tau_0 \leq \tau < \omega_0$,

$$|x_\lambda^{\mu_{\nu\tau}}| \leq ||x^{\mu_{\nu\tau}}||_{R_\nu} \leq ||x^{\mu_{\nu\tau}}||_{R_{\mu_{\nu\tau}}} = \varrho_{\mu_{\nu\tau}} \leq \varrho_\mu,$$

and $\tau \to \omega_0$ yields $|y_\lambda^\nu| \leq \varrho_\mu$, where y^ν is defined by (6). Hence $\varrho_\nu \leq ||y^\nu||_{R_\nu} \leq \varrho_\mu$, and (8) follows. Clearly,

$$N_\mu N_\nu = 0 \qquad (\mu < \nu < \omega_1).$$

Hence there exists an ordinal $n < \omega_1$ such that

$$n = \min\,[N_\nu = 0]\,\nu.$$

I can now prove that

$$\varrho_\mu > \varrho_\nu \qquad (\mu < \nu < n). \tag{9}$$

Assume, to the contrary, that $\mu < \nu < n$; $\varrho_\mu = \varrho_\nu$. Then

$$\varrho_\mu \geq \varrho_{\mu+1} \geq \varrho_\nu = \varrho_\mu; \qquad \varrho_\mu = \varrho_{\mu+1}.$$

Put $x' = \frac{1}{2}(x^\mu + x^{\mu+1})$. Then $x' \in Z_{\mu+1}$. By definition of n, I can choose a number λ in $N_{\mu+1}$. Then

$$|x'_\lambda| \leq \tfrac{1}{2}|x_\lambda^\mu| + \tfrac{1}{2}|x_\lambda^{\mu+1}| < \tfrac{1}{2}\varrho_\mu + \tfrac{1}{2}\varrho_{\mu+1} = \varrho_{\mu+1}. \tag{10}$$

Here I have used (7) and the fact that $\lambda \in R_\mu - N_\mu$. But (10) contradicts the definition of $N_{\mu+1}$. Hence (9) follows.

I finally show that x^n is a minimal sequence of X. Put $z = x^n$. Then $||z||_{R_n} = \varrho_n \leq \varrho_\nu$ $(\nu < n)$. Also,

$$|z_\lambda| = |x_\lambda^n| < \varrho_n \qquad (\lambda \in R_n),$$

in view of (7) and the fact that $N_n = 0$. Furthermore,

$$\{-|z_\lambda| : \lambda < \omega_0\} = A + B,$$

where

$$A = \{-|z_\lambda| : \lambda \in L_n\} = \{-\varrho_\nu : \nu < n\}; \qquad B = \{-|z_\lambda| : \lambda \in R_n\}.$$

A is well-ordered, and B has no least element. Every element of A is less than every element of B, and

$$N_\nu = \{\lambda : |z_\lambda| = \varrho_\nu\} \qquad (\nu < n).$$

All this shows that, in the notation introduced at the end of § 3, $n = n(z)$; $N_\nu = N_\nu(z)$; $\varrho_\nu = \varrho_\nu(z)$.

Now let $x \in X$. There are the following cases.

Case 1. $x = z$ in L_n.

Case 11. $x \overline{\in} Y$. Then $||x|| > c$; $||x||_{R_n} > c \geq ||z||_{R_n} = \varrho(z)$.

Case 12. $x \in Y$. Then $x \in Y_n$; $||x||_{R_n} \geq \varrho_n = \varrho(z)$.

Case 2. There is $\nu < n$; $\lambda \in N_\nu$ such that $x = z$ in L_ν, and $x_\lambda \neq z_\lambda$.

Case 21. $x \overline{\in} Y$. Then $||x||_{R_\nu} > c \geq ||z||_{R_\nu} = \varrho_\nu(z)$.

Case 22. $x \in Y$.

Case 221. $||x||_{R_\nu} > \varrho_\nu$. Then $||x||_{R_\nu} > \varrho_\nu = \varrho_\nu(z)$.

Case 222. $||x||_{R_\nu} \leq \varrho_\nu$. Then $x \in Z_\nu$. Hence, if $x' = \frac{1}{2}(x+z)$, then $x' \in Z_\nu$; $|x_\lambda| = |z_\lambda| = \varrho_\nu$; $x_\lambda \neq z_\lambda$, and therefore $|x'_\lambda| = |\frac{1}{2}(x_\lambda + z_\lambda)| < \varrho_\nu$. But this contradicts the definition of N_ν.

This proves that z is a minimal sequence of X and so completes the proof of Theorem 3.

8. I now prove Theorem 2. Suppose that the hypotheses of Theorem 2 are satisfied. Let $R k A = p$.

It is possible, by a standard process, to replace the system of equations defining X by an equivalent system

$$\sum [\lambda < \omega_0]\, b_{\rho\lambda}\, x_\lambda + b_\rho = 0 \qquad (\rho < m)$$

which is such that $\sum [\lambda < \omega_0] |b_{\rho\lambda}| < \infty$ $(\rho < m)$,

$$b_{\rho\lambda} = 0 \qquad (p \leq \rho < m;\ \lambda < \omega_0).$$

It then follows from $X \neq 0$ that $b_\rho = 0$ $(p \leq \rho < m)$. Hence there is $y \in X$ such that $y_\lambda = 0$ $(\lambda_0 \leq \lambda < \omega_0)$, for some $\lambda_0 < \omega_0$. Now, if $||y|| = c$, we have $c < \infty$, and the set $Y = \{x : x \in X;\ ||x|| \leq c\}$ is not empty. Also, if $x', x'' \in Y$, then $\frac{1}{2}(x' + x'') \in Y$, so that Y is convex. Y is weakly closed. For, let $x^s \in Y$ $(s < \omega_0)$,

$$\lim [s \to \omega_0]\, x^s_\lambda = x^*_\lambda \qquad (\lambda < \omega_0).$$

Then $|x^s_\lambda| \leq c$; $|x^*_\lambda| \leq c$ and, by uniform convergence,

$$\sum a_{\rho\lambda}\, x^*_\lambda = \lim [s \to \omega_0] \sum a_{\rho\lambda}\, x^s_\lambda = -a_\rho \qquad (\rho < m).$$

Theorem 3 can be applied to X and yields a minimal sequence z of X. Let $n(z) = n$; $\varrho_\nu(z) = \varrho_\nu$, etc.

If $M, N \subset [0, \omega_0]$, then $[a_{\rho\lambda}]_{\rho \in M;\, \lambda \in N}$ denotes the minor of the matrix $A = [a_{\rho\lambda}]$ obtained by restricting ρ to M and λ to N. If M and N are of the

same finite cardinal, then $|a_{r\lambda}|_{r\in M;\,\lambda\in N}$ is the determinant of that minor. The conditions on r and λ may also take forms different from $r\in M;\ \lambda\in N$.

Put

$$p_\nu = R\,k\,[a_{r\lambda}]_{r<m;\,\lambda\in R_\nu}.$$

In particular, $p_\nu = 0$ if $R_\nu = 0$. Then $p_\nu \geq p_{\nu+1}$ $(\nu<\omega_1)$. Let $\nu<\omega_1$; $p_\nu = p_{\nu+1} = q>0$. I want to deduce a contradiction. I may assume that $[0,q]\subset R_{\nu+1}$ and $|a_{r\lambda}|_{r,\lambda<q}\neq 0$. Then there is $x'\in X$ such that $x'=z$ in L_ν; $x'=0$ in $R_\nu[q,\omega_0]$. Put, for $s<\omega_0$,

$$z^s = (1-2^{-s})\,z + 2^{-s}\,x'.$$

Then $z^s\in X$; $z^s = z$ in L_ν. Hence, by definition of ϱ_ν, $\|z^s\|_{R_\nu}\geq\varrho_\nu$. Let $\lambda\in R_\nu[q,\omega_0]$. Then

$$|z^s_\lambda| = (1-2^{-s})\,|z_\lambda| \leq (1-2^{-s})\,\varrho_\nu.$$

If, first of all, $\varrho_\nu>0$, then it follows that

$$\|z^s\|_{R_\nu[q,\,\omega_0]} \leq (1-2^{-s})\,\varrho_\nu < \varrho_\nu$$

and therefore

$$\|z^s\|_{R_\nu[0,q]} \geq \varrho_\nu.$$

This also holds if $\varrho_\nu = 0$. By making $s\to\omega_0$ we obtain

$$\|z\|_{R_\nu[0,q]} \geq \varrho_\nu.$$

Hence there is $\lambda\in R_\nu[0,q]$ such that $|z_\lambda|\geq\varrho_\nu$. Then, by definition of N_ν,

$$\lambda\in N_\nu[0,q]\subset L_{\nu+1}\,R_{\nu+1} = 0$$

which is the desired contradiction. This proves that, for $\nu<\omega_1$, either $p_\nu > p_{\nu+1}$ or $p_\nu = p_{\nu+1} = 0$. Therefore

$$p_0 > p_1 > \cdots > p_k = 0,$$

for some $k\leq p_0$. Then $a_{r\lambda}=0$ $(r<m;\ \lambda\in R_k)$, and X is defined by the system of equations

$$\sum[\lambda\in L_k]\,a_{r\lambda}\,x_\lambda + a_r = 0 \qquad (r<m).$$

Then $z'\in X$, where $z'=z$ in L_k; $z'=0$ in R_k. Then, in the notation used in the proof of Theorem 2, $z'\in Y_k$ and, by definition of ϱ_k, $0=\|z'\|_{R_k}\geq\varrho_k$. If now, $k+1<n$, then $\varrho_k>\varrho_{k+1}\geq 0$ which is false. Hence

$$n\leq k+1\leq p_0+1 = 1+R\,k\,A.$$

In order to complete the proof of Theorem 2 I need only show that $R_n = 0$.

We have $R_n = R_{n+1}$; $p_n = p_{n+1} = 0$; $a_{r\lambda}=0$ $(r<m;\ \lambda\in R_n)$. Let $z''=z$ in L_n; $z''=0$ in R_n. Then $z''\in Y_n$; $0=\|z''\|_{R_n}\geq\varrho_n$. If, now, $R_n\neq 0$ then we could choose an index λ in R_n. Then $\lambda\in R_n = R_n - N_n$ and therefore $|z_\lambda|<\varrho_n = 0$ which is a contradiction. Hence $R_n = 0$, and Theorem 2 is proved.

9. I conclude this note with some illustrations and some remarks on further extensions.

(i) Let $0 \neq N_0 < [0, \omega_0]$; $N_1 = [0, \omega_0] - N_0 \neq 0$; $\varrho_0 > \varrho_1 = 0$. Then the following set X, defined by one linear equation, has $n = 2$ and the given N_ν, ϱ_ν associated with it in the sense of Theorems 2 and 3. Choose numbers $a_{0\lambda} > 0$ $(\lambda \in N_0)$ such that $\varrho_0 \sum [\lambda \in N_0] a_{0\lambda} = 1$, and put $a_{0\lambda} = 0$ $(\lambda \in N_1)$. Let X be the set of all sequences x such that

$$\sum [\lambda < \omega_0] a_{0\lambda} x_\lambda - 1 = 0,$$

where the series is stipulated be to be absolutely convergent. Let z be defined by $z_\lambda = \varrho_\nu$ $(\nu < 2;\ \lambda \in N_\nu)$. Then $z \in X$; $||z|| = \varrho_0$. If $||x|| < \varrho_0$, then

$$|\sum a_{0\lambda} x_\lambda| < 1; \quad x \overline{\in} X.$$

If $||x|| = \varrho_0$; $\lambda_0 \in N_0$; $x_{\lambda_0} \neq \varrho_0$, then $x \overline{\in} X$. This proves that z is the unique minimal sequence of X, according to Theorem 2, and that $n(z) = 2$, $N_\nu(z) = N_\nu$; $\varrho_\nu(z) = \varrho_\nu$ $(\nu < 2)$.

(ii) Let $0 < n < \omega_1$; $0 \neq N_\nu < [0, \omega_0)$ $(\nu < n)$; $\varrho_\nu > \varrho > 0$ $(\nu < n)$;

$$\varrho_\mu > \varrho_\nu; \quad N_\mu N_\nu = 0 \qquad (\mu < \nu < n).$$

Suppose that the set $[0, \omega_0] - \sum [\nu < n] N_\nu = R$ is infinite. For any given n in the range $0 < n < \omega_1$ it is possible to choose N_ν and ϱ_ν, ϱ such that these conditions hold. I now construct a sequence z such that $\varrho(z) = \varrho$; $n(z) = n$;

$$N_\nu(z) = N_\nu; \quad \varrho_\nu(z) = \varrho_\nu \qquad (\nu < n).$$

Put

$$z_\lambda = \varrho_\nu \qquad (\nu < n;\ \lambda \in N_\nu);$$
$$z_\lambda = \varrho\, \lambda (\lambda + 1)^{-1} \qquad (\lambda \in R).$$

Then z has the required properties.

(iii) An examination of the argument used to prove Theorem 3 shows that all that is required of the set of all complex numbers is that it is a vector space over the field of real numbers possessing a metric under which it is closed and for which every closed sphere is strictly convex. It follows that in Theorem 3 the sequences of complex numbers may be replaced by sequences $(x_0, x_1, \ldots)$ such that, for $\lambda < \omega_0$, x_λ is any element of a given vector space S_λ of the type just described.

(iv) The conclusion of Theorem 3 no longer holds if the sequences x are not of type ω_0 but of a non-denumerable well-order type. This is seen by considering the following example. Let X be the set of all sequences

$$x = (x_0, x_1, \ldots, x_{\omega_\nu}, \ldots, x_\lambda, \ldots) \qquad (\lambda < \omega_1),$$

of type ω_1, which have that the property that, for $\lambda < \omega_1$, x_λ is a real number in the range $0 \leq x_\lambda \leq 1$, and there exists an ordinal $l = l(x) < \omega_1$ such that $x_\lambda = 1$ $(l \leq \lambda < \omega_1)$. Then the set X is convex and weakly closed but does not

possess any minimal sequence, i.e. a sequence z with properties analogous to those described in the case of sequences of type ω_0. For, we note that, in our case,

$$\varrho_0 = \inf [x \in X] \, ||x|| = 1 .$$

Let $z \in X$. Then there is $\lambda < \omega_1$ such that $z_\lambda = 1$. If, now, x is such that $x_\lambda = 0$, while $x = z$ in $[0, \omega_1] - \{\lambda\}$, then $x \in X$; $||x|| = \varrho_0$, but $x_\lambda \neq z_\lambda = \varrho_0$.

This shows that z is not a minimal sequence of X.

Obviously, when dealing with ω_1-sequences it is necessary to modify the definition of weak closure of a set by taking into consideration weak limits of sequences x^s $(s < \omega_1)$. In this, stronger, sense, the set just defined would not be weakly closed. But even with this stronger definition of closure there does not appear to be an obvious way of extending Theorem 3 to ω_1-sequences. This is due to the fact that Cantor's diagonal process played a prominent part in the proof of Theorem 3, and there does not seem to exist an equally effective analogue of this process in the case when ω_0-sequences are replaced by ω_1-sequences.

Reference

[1] Rogosinski, W. W.: On finite systems of linear equations with an infinity of unknowns. Math. Z. **63**, 97—108 (1955).

The University, Reading (England)

(Eingegangen am 9. Juli 1955)

FENCHEL, W.
Math. Zeitschr. Bd. 63, S. 496—506 (1956)

Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten

ISSAI SCHUR zum Gedächtnis

Von

WERNER FENCHEL

Das Problem, das im folgenden diskutiert wird, läßt sich für Funktionen von zwei Variablen x_1 und x_2 folgendermaßen formulieren: In einem konvexen Bereich C der x_1x_2-Ebene sei eine Schar (geschlossener oder offener) konvexer Kurven gegeben, die C einfach und lückenlos überdecken. Gefragt wird, unter welchen Bedingungen eine stetige konvexe Funktion $f(x_1, x_2)$ existiert, deren Niveaukurven diese Kurven sind. Ist die Kurvenschar durch eine Gleichung der Form $\varphi(x_1, x_2) = \text{konst.}$ gegeben, wo $\varphi(x_1, x_2)$ eine in C stetige Funktion ist, so läuft dies auf die Aufstellung notwendiger und hinreichender Bedingungen hinaus, die diese Funktion erfüllen muß, wenn es eine stetige monotone Funktion $F(t)$ einer Variablen derart geben soll, daß die Funktion $f(x_1, x_2) = F(\varphi(x_1, x_2))$ konvex ist.

Diese Frage, die auf den ersten Blick trivial erscheinen mag, ist erstmalig von B. DE FINETTI [3] aufgeworfen und für einen beschränkten abgeschlossenen Bereich C und eine Schar geschlossener Kurven eingehend behandelt worden. An anschaulichen Beispielen zeigt DE FINETTI, daß die Schar konvexer Kurven keineswegs willkürlich vorgeschrieben werden kann, sondern verschiedenartigen Einschränkungen unterworfen ist, die sowohl ihr lokales als auch ihr globales Verhalten betreffen. Notwendige und hinreichende Bedingungen für die Lösbarkeit der gestellten Aufgabe sind implizit in der Forderung enthalten, daß eines der beiden von DE FINETTI angegebenen Verfahren zur Konstruktion der gesuchten konvexen Funktion zum Ziele führen soll. In einfacher und durchsichtiger Form dürften sich solche Bedingungen kaum formulieren lassen. Es liegt daher nahe zu versuchen, eine etwas bessere Einsicht in ihre Natur dadurch zu gewinnen, daß man die Frage unter geeigneten Differenzierbarkeitsvoraussetzungen behandelt. In der Arbeit DE FINETTIs wird dieser Fall kurz und auf eine Weise erwähnt, die den Eindruck erwecken könnte, daß dann die Frage eine ganz einfache Antwort zulasse ([3] S. 173—174). Dies ist aber keineswegs der Fall, nicht einmal, wenn man die Kurvenschar analytisch annimmt. Es ist daher vielleicht nicht überflüssig, nochmals auf den Gegenstand zurückzukommen.

Im folgenden wird die Fragestellung verallgemeinert, indem Funktionen von beliebig vielen Variablen und ein nicht notwendig beschränkter konvexer Definitionsbereich in Betracht gezogen werden. Andererseits wird vorausgesetzt, daß die gegebene Funktion $\varphi(x_1, \ldots, x_n)$ und die gesuchte $F(t)$,

für die also $F(\varphi(x_1, \ldots, x_n))$ konvex sein soll, zweimal differenzierbar sind. Es werden notwendige und hinreichende Bedingungen für die Existenz einer solchen Funktion $F(t)$ aufgestellt, die das Verhalten von φ teils in den einzelnen Punkten, teils im großen betreffen. Die lokalen Bedingungen (§ 2) bestehen in Ungleichungen und algebraischen Relationen zwischen den ersten und zweiten partiellen Ableitungen von φ und lassen einfache geometrische Deutungen zu. Im Fall von zwei Variablen ergibt sich beispielsweise, daß an einer Stelle, an der die Krümmung der betreffenden Niveaukurve $\varphi(x_1, x_2) = \text{konst.}$ verschwindet, auch die Gausssche Krümmung der Fläche $z = \varphi(x_1, x_2)$ verschwinden muß. Die globalen Bedingungen (§ 3) sind dagegen auch in dem hier betrachteten Fall weniger durchsichtig, und eine weitere Aufspaltung wäre nicht ohne Interesse. Sie enthalten beispielsweise implizit die einfache notwendige Bedingung, daß alle Niveaumannigfaltigkeiten $\varphi(x_1, \ldots, x_n) = \text{konst.}$ dieselben unendlich fernen Punkte haben müssen (§ 5).

Der Verfasser hat sich mit dem de Finettischen Problem auf Anregung von A. W. Tucker beschäftigt und über die genannten Resultate in einer Vorlesungsreihe über konvexe Mengen und Funktionen [2] berichtet, die er 1951 im Rahmen eines vom Office of Naval Research unterstützten Forschungsprogramms an der Universität Princeton gehalten hat.

1. Vorbemerkungen. Es werden reelle Funktionen betrachtet, die in einem Bereich des n-dimensionalen euklidischen Raumes E^n definiert sind. (Die zu behandelnde Aufgabe hat zwar affininvarianten Charakter, es ist aber vielfach bequem, metrische Begriffe als Hilfsmittel heranzuziehen.) Im E^n wird ein rechtwinkliges Koordinatensystem zugrunde gelegt. Der Punkt mit den Koordinaten $x_1, \ldots, x_n$ wird im allgemeinen kurz mit x bezeichnet werden. Sind $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ zwei Punkte und λ eine reelle Zahl, so werden λx und $x + y$ wie üblich durch

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n), \qquad x + y = (x_1 + y_1, \ldots, x_n + y_n)$$

definiert.

Es sei nun C ein (nicht notwendig beschränkter) konvexer Bereich im E^n, der, um unwesentliche Komplikationen zu vermeiden, stets als offen vorausgesetzt wird. Eine in C definierte Funktion $f(x)$ heißt (nach unten) konvex, wenn für beliebige Punkte x und y aus C und $0 \leq \Theta \leq 1$ die Ungleichung

$$f((1 - \Theta) x + \Theta y) \leq (1 - \Theta) f(x) + \Theta f(y)$$

gilt. Dies ist gleichbedeutend damit, daß die Menge derjenigen Punkte $(x, z) = (x_1, \ldots, x_n, z)$ des $(n+1)$-dimensionalen Raumes E^{n+1}, für welche $x \in C$ und $z \geq f(x)$ gilt, konvex ist. Von den wohlbekannten Eigenschaften der in C nicht konstanten konvexen Funktionen kommen die folgenden zur Anwendung:

a) Eine in C konvexe Funktion $f(x)$ ist dort stetig.

b) Eine in C konvexe Funktion $f(x)$ hat entweder überhaupt keine Extrema oder nur ein absolutes Minimum, das in einer relativ zu C abgeschlossenen konvexen Teilmenge von C angenommen wird.

c) Ist $f(x)$ konvex in C, und ist t eine Zahl aus dem Wertebereich von f, so bilden die Punkte x, in denen $f(x) \leq t$ gilt, eine relativ zu C abgeschlossene konvexe Teilmenge von C.

d) Existieren die partiellen Ableitungen $f_i = \partial f / \partial x_i$, $i = 1, \ldots, n$, einer in C konvexen Funktion $f(x)$ überall in C, so sind sie stetig, so daß f also insbesondere differenzierbar ist, und sie verschwinden gleichzeitig nur in den eventuellen Minimumspunkten von f.

e) Ist die Funktion $f(x)$ zweimal differenzierbar in C, sind also sie selbst und ihre ersten Ableitungen differenzierbar, so ist sie dann und nur dann konvex, wenn die quadratische Form

$$\sum_{i,j=1}^{n} f_{ij}(x)\, \xi_i \xi_j, \qquad f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j},$$

in den Variablen $\xi_1, \ldots, \xi_n$ für jeden Punkt x aus C positiv semidefinit ist.

2. Lokale notwendige Bedingungen. Wir formulieren nun die zu lösende Aufgabe. Gegeben sei eine im offenen konvexen Bereich C zweimal differenzierbare, nicht konstante Funktion $\varphi(x)$. Es werde

$$\alpha = \inf_{x \in C} \varphi(x), \qquad \beta = \sup_{x \in C} \varphi(x)$$

gesetzt, wo die Werte $-\infty$ und $+\infty$ von α bzw. β zugelassen sind. Der Wertevorrat von $\varphi(x)$ ist dann ein Intervall mit den Endpunkten α und β. Gesucht ist eine in diesem Intervall monoton wachsende, zweimal differenzierbare Funktion $F(t)$, für welche $f(x) = F(\varphi(x))$ konvex in C ist.

Wir beginnen mit der Herleitung notwendiger Bedingungen. Es sei also $f(x) = F(\varphi(x))$ konvex in C. Mit N_α werde die (eventuell leere) Teilmenge von C bezeichnet, in der $\varphi(x) = \alpha$ ist. Mit den Bezeichnungen

$$F' = \frac{dF}{dt}, \qquad \varphi_i = \frac{\partial \varphi}{\partial x_i}$$

folgt aus

$$f_i = F' \varphi_i, \qquad i = 1, \ldots, n, \tag{1}$$

und den Eigenschaften b) und c) von f, daß die Ableitungen φ_i von φ außerhalb N_α nicht gleichzeitig verschwinden können. Wir haben also

$$k^2 = \varphi_1^2 + \cdots + \varphi_n^2 > 0 \quad \text{für} \quad x \notin N_\alpha. \tag{2}$$

Wir formulieren dies als Bedingung

I. *Die Funktion $\varphi(x)$ hat entweder keine stationären Werte oder nur ein absolutes Minimum.*

Speziell folgt, daß φ die obere Grenze β nicht annehmen kann, so daß der Wertebereich $\alpha \leq t < \beta$ geschrieben werden kann, wo die Klammern um das Gleichheitszeichen andeuten, daß es dann und nur dann in Frage kommt, wenn φ ein Minimum hat.

Im Hinblick auf das folgende bemerken wir noch, daß aus d), (1) und $F' \gtrless 0$

$$F'(t) > 0 \qquad \text{für} \qquad \alpha < t < \beta \tag{3}$$

folgt.

Durch Differentiation von (1) erhält man mit den obigen analogen Bezeichnungen

$$f_{ij} = F' \varphi_{ij} + F'' \varphi_i \varphi_j, \qquad i,j = 1, \ldots, n,$$

und aus e) folgt daher, daß die quadratische Form

$$F' \sum_{i,j=1}^{n} \varphi_{ij} \xi_i \xi_j + F'' \left(\sum_{i=1}^{n} \varphi_i \xi_i \right)^2$$

an jeder Stelle x von C positiv semidefinit sein muß. Ist N_α nicht leer, so ist dies wegen $F'(\alpha) \gtrless 0$ in den Punkten x von N_α stets von selbst erfüllt; denn da diese Minimumstellen von φ sind, gilt $\varphi_i = 0$, $i = 1, \ldots, n$, und die Form $\sum \varphi_{ij} \xi_i \xi_j$ ist positiv semidefinit. Wir nehmen daher im folgenden an, daß x nicht zu N_α gehört, daß also $\varphi(x) > \alpha$ ist. Dann muß wegen (3) für jedes solche x die quadratische Form

$$Q_\sigma(\xi, \xi) = \sum_{i,j=1}^{n} \varphi_{ij} \xi_i \xi_j + \sigma \left(\sum_{i=1}^{n} \varphi_i \xi_i \right)^2 \tag{4}$$

für ein gewisses $\sigma = \sigma(x)$, nämlich $\sigma = F''/F'$, positiv semidefinit sein. Als unmittelbare Folge ergibt sich:

II. *Für jedes x aus C ist die Form $\sum \varphi_{ij} \xi_i \xi_j$ unter der Nebenbedingung $\sum \varphi_i \xi_i = 0$ positiv semidefinit.*

Für die Punkte x von N_α ist dies nämlich nach einer soeben gemachten Bemerkung eine Folge von I.

Daß II notwendig ist, könnte man natürlich auch daraus schließen, daß die Menge der Punkte x, in denen $\varphi(x) \leqq t$ ist, für jedes t, $\alpha (\leqq) t < \beta$, wegen § 1, c) konvex sein muß, da sie ja mit der Menge der Punkte, in denen die konvexe Funktion $F(\varphi(x)) \leqq F(t)$ ist, übereinstimmt. Es sei bemerkt, daß die Konvexität dieser Mengen, wie sehr leicht zu sehen, damit gleichbedeutend ist, daß die Funktion φ „quasikonvex" ist. Dies soll besagen, daß sie auf jeder in C enthaltenen abgeschlossenen Strecke ihr Maximum in einem der Endpunkte annimmt, daß also für alle x und y aus C und $0 \leqq \Theta \leqq 1$

$$\varphi\big((1-\Theta)\, x + \Theta y\big) \leqq \operatorname{Max}\left[\varphi(x), \varphi(y)\right].$$

Wir haben nun weiter die Bedingung dafür aufzustellen, daß es zu jedem festen, nicht in N_α enthaltenen x eine Zahl σ derart gibt, daß die Form $Q_\sigma(\xi, \xi)$ für alle ξ_i mit $\sum \varphi_i \xi_i \neq 0$ nichtnegativ ist, oder anders ausgedrückt, daß

$$\sup_{\sum \varphi_i \xi_i \neq 0} \left(- \frac{\sum \varphi_{ij} \xi_i \xi_j}{(\sum \varphi_i \xi_i)^2} \right) < \infty \tag{5}$$

ist. Zu diesem Zweck betrachten wir das charakteristische Polynom $\Gamma_{Q_\sigma}(\lambda)$ der Form Q_σ. Mit der üblichen Bedeutung von δ_{ij} und der Festsetzung, daß

die Indizes i und j in den folgenden Determinantenformeln die Werte $1, \ldots, n$ durchlaufen sollen, haben wir

$$(6)\quad \begin{cases} \Gamma_{Q_\sigma}(\lambda) = |\varphi_{ij} - \lambda\,\delta_{ij} + \sigma\,\varphi_i\,\varphi_j| \\ \qquad = \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} + \sigma\,\varphi_i\,\varphi_j & \varphi_i \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ -\sigma\,\varphi_j & 1 \end{vmatrix} \\ \qquad = |\varphi_{ij} - \lambda\,\delta_{ij}| - \sigma \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix}. \end{cases}$$

Wir setzen zur Abkürzung $P(\xi, \xi) = \sum \varphi_{ij}\,\xi_i\,\xi_j$ und bezeichnen mit $P^*(\xi, \xi)$ dieselbe Form, wenn die Variablen ξ_i der Bedingung $\sum \varphi_i\,\xi_i = 0$ unterworfen sind. Die charakteristischen Polynome dieser Formen sind

$$\Gamma_P(\lambda) = |\varphi_{ij} - \lambda\,\delta_{ij}|$$

und

$$\Gamma_{P^*}(\lambda) = -\frac{1}{k^2} \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix}.$$

Die Richtigkeit der letzteren Gleichung kann man z.B. so einsehen: Beide Seiten sind invariant bei orthogonalen Transformationen der Koordinaten x_i, da solche Transformationen auf orthogonale Transformationen der Variablen ξ_i hinauslaufen (vgl. z.B. [4] S. 241–242). Wählt man nun das Koordinatensystem so, daß an der betreffenden Stelle $\varphi_1 = \cdots = \varphi_{n-1} = 0$, $\varphi_n = k$ wird, so vereinfacht sich die Nebenbedingung zu $\xi_n = 0$, und die rechte Seite geht tatsächlich in das charakteristische Polynom von $\sum_{i,j=1}^{n-1} \varphi_{ij}\,\xi_i\,\xi_j$ über. Hiernach kann (6) also

$$\Gamma_{Q_\sigma}(\lambda) = \Gamma_P(\lambda) + \sigma\,k^2\,\Gamma_{P^*}(\lambda)$$

geschrieben werden. Bezeichnen wir die ν-ten elementarsymmetrischen Funktionen der Eigenwerte von P, P^* und Q_σ mit S_ν, S_ν^* bzw. $T_\nu^{(\sigma)}$ und setzen $S_0 = S_0^* = T_0^{(\sigma)} = 1$, so haben wir

$$\Gamma_P(\lambda) = S_n - S_{n-1}\,\lambda + \cdots + (-1)^n\,S_0\,\lambda^n,$$

$$\Gamma_{P^*}(\lambda) = S_{n-1}^* - S_{n-2}^*\,\lambda + \cdots + (-1)^{n-1}\,S_0^*\,\lambda^{n-1},$$

$$\Gamma_{Q_\sigma}(\lambda) = T_n^{(\sigma)} - T_{n-1}^{(\sigma)}\,\lambda + \cdots + (-1)^n\,T_0^{(\sigma)}\,\lambda^n.$$

Die Forderung, daß Q_σ positiv semidefinit sein soll, ist also gleichbedeutend mit

$$(7)\qquad T_\nu^{(\sigma)} = S_\nu + \sigma\,k^2\,S_{\nu-1}^* \geq 0, \qquad \nu = 1, \ldots, n,$$

wo das Verschwinden eines $T_\nu^{(\sigma)}$ das aller folgenden nach sich zieht.

Es sei nun $r-1$ der (natürlich im allgemeinen von x abhängige) Rang von P^*. Dann haben wir nach II

$$(8)\qquad S_\nu^* > 0 \quad \text{für} \quad 0 \leq \nu \leq r-1, \qquad S_\nu^* = 0 \quad \text{für} \quad r \leq \nu \leq n-1,$$

so daß aus (7) speziell

$$(9)\qquad S_{r+1} \geq 0$$

folgt. Zwischen den Eigenwerten

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

und

$$\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_{n-1}^*$$

von P bzw. P^* bestehen infolge ihrer Minimum-Maximum-Eigenschaften bekanntlich die Ungleichungen

$$\lambda_1 \geq \lambda_1^* \geq \lambda_2 \geq \cdots \geq \lambda_\nu \geq \lambda_\nu^* \geq \cdots \geq \lambda_{n-1}^* \geq \lambda_n$$

(vgl. z.B. [1] S. 28). Wegen (8) ist nun

$$\lambda_\nu^* > 0 \quad \text{für} \quad 1 \leq \nu \leq r-1, \qquad \lambda_\nu^* = 0 \quad \text{für} \quad r \leq \nu \leq n-1,$$

folglich

$$\lambda_\nu > 0 \quad \text{für} \quad 1 \leq \nu \leq r-1$$

und, falls $r < n$ ist,

$$\lambda_\nu \geq 0, \quad \lambda_\nu = 0 \quad \text{für} \quad r+1 \leq \nu \leq n-1, \quad \lambda_n \leq 0,$$

insbesondere also $S_{r+1} = \lambda_1 \ldots \lambda_r \lambda_n \leq 0$. Zusammen mit (9) gibt dies

$$S_{r+1} = 0 \quad \text{falls} \quad r < n;$$

d.h. der Rang von P kann höchstens gleich r sein. Wir formulieren dies als weitere notwendige Bedingung:

III. *Hat im Punkt $x \in C$ die Form $\sum \varphi_{ij}\, \xi_i \xi_j$ unter der Nebenbedingung $\sum \varphi_i \xi_i = 0$ den Rang $r(x)-1$, so hat sie ohne diese Bedingung höchstens den Rang $r(x)$.*

Dies wurde zwar nur für die nicht zu N_α gehörigen Punkte von C bewiesen; für $x \in N_\alpha$ ist es aber wegen $\varphi_i(x) = 0$, $i = 1, \ldots, n$, von selbst erfüllt.

Die Notwendigkeit einer Bedingung von der Art III ist auch anschaulich unmittelbar einleuchtend. Für $n = 2$ betrachte man in der $x_1 x_2$-Ebene etwa das Geradenbüschel mit dem Nullpunkt als Scheitel. In einem in der Halbebene $x_1 > 0$ enthaltenen konvexen Gebiet C sind die Büschelgeraden die Niveaukurven der Funktion $\varphi = x_2/x_1$. Die Bedingungen I und II sind offensichtlich erfüllt. Für jede wachsende Funktion $z = F(t)$ ist aber die Fläche $z = F(x_2/x_1)$ eine windschiefe Regelfläche und daher niemals konvex. Daß III auch bei Konvexität im engeren Sinne der Niveaukurven nicht überall erfüllt zu sein braucht, zeigt das Beispiel der Funktion $\varphi = (x_2 - x_1^4)/(x_1 + 1)$ in der Halbebene $x_1 > -1$. Auch hier sind I und II erfüllt, während III für $x_1 = 0$ nicht erfüllt ist. Es beruht dies darauf, daß die Niveaukurven $x_2 = x_1^4 + t(x_1 + 1)$ in ihren Schnittpunkten mit der x_2-Achse verschwindende Krümmung, aber nicht parallele Tangenten haben.

Im allgemeinen Fall läßt III folgende geometrische Deutung zu: Verschwinden in einem Punkt der Hyperfläche $\varphi(x) = \text{konst.}$ des E^n genau $n-r$ ihrer Hauptkrümmungen, so müssen an dieser Stelle auch mindestens $n-r$ der Hauptkrümmungen der Hyperfläche $z = \varphi(x)$ des E^{n+1} verschwinden.

3. Globale notwendige Bedingungen. Die Zahlen σ, für welche $Q_\sigma(\xi, \xi)$ positiv semidefinit ist, müssen nach (7) der Ungleichung

$$\sigma \geq \sigma_0 = \max_{1 \leq \nu \leq r} \left(- \frac{S_\nu}{k^2 S^*_{\nu-1}} \right)$$

genügen. Da nun, wie schon bemerkt, das Verschwinden eines $T_\nu^{(\sigma)}$ das aller folgenden nach sich zieht, muß das Maximum für $\nu = r$ erreicht werden, so daß

$$\sigma_0 = - \frac{S_r}{k^2 S^*_{r-1}}.$$

Wie aus (4) hervorgeht, ist σ_0 zugleich die in (5) auftretende obere Grenze.

Wir ziehen nun die Abhängigkeit vom Punkt x in Betracht. Soll es eine Funktion $F(t)$ mit den verlangten Eigenschaften geben, so muß nach (4)

$$\sigma_0(x) \leq \frac{F''(\varphi(x))}{F'(\varphi(x))}$$

gelten. Da die rechte Seite auf jeder Niveaumannigfaltigkeit konstant ist, ergibt sich als weitere notwendige Bedingung:

IV. *Für jedes feste t mit $\alpha < t < \beta$ ist*

$$G(t) = \sup_{\varphi(x) = t} \left(- \frac{S_r}{k^2 S^*_{r-1}} \right) < \infty.$$

Schließlich verlangt die Aufgabe, daß eine in $\alpha \leqq t < \beta$ zweimal differenzierbare Funktion $F(t)$ mit $F'(t) > 0$ für $\alpha < t < \beta$ existiert, so daß für diese t

$$\frac{F''(t)}{F'(t)} \geq G(t).$$

Ob sich diese Bedingung wesentlich reduzieren läßt, muß dahingestellt bleiben. Da Stetigkeit der zweiten Ableitungen von φ nicht vorausgesetzt ist und die Niveaumannigfaltigkeiten sich ins Unendliche erstrecken können, wird man im allgemeinen von $G(t)$ wohl kaum mehr als die Meßbarkeit behaupten können. Wir formulieren daher als notwendige Bedingung:

V. *Die Funktion $G(t)$, $\alpha < t < \beta$, hat eine Majorante, die die logarithmische Ableitung einer in $\alpha < t < \beta$ positiven und in $\alpha \leqq t < \beta$ differenzierbaren Funktion ist.*

4. Das Hinreichen der Bedingungen. Der Beweis dafür, daß die aufgestellten Bedingungen I bis V auch hinreichend sind, ist sehr einfach. Es sei eine in einem offenen konvexen Bereich C des E^n definierte, zweimal differenzierbare Funktion $\varphi(x)$ gegeben, die I bis V erfüllt. Nach V gibt es eine Funktion $H(t)$, die in $\alpha \leqq t < \beta$ differenzierbar und in $\alpha < t < \beta$ positiv ist, derart daß

$$\frac{H'(t)}{H(t)} \geq G(t), \qquad \alpha < t < \beta.$$

Es sei $F(t)$ ein unbestimmtes Integral von $H(t)$. Dann ist $F(t)$ zweimal differenzierbar und monoton wachsend in $\alpha \leqq t < \beta$. Wir haben zu zeigen, daß

die quadratische Form

$$F'(\varphi(x))\sum \varphi_{ij}(x)\,\xi_i\,\xi_j + F''(\varphi(x))\,(\sum \varphi_i\,\xi_i)^2 \tag{10}$$

für jedes $x \in C$ positiv semidefinit ist.

Besitzt $\varphi(x)$ gemäß I ein Minimum α, so ist diese Behauptung offenbar für die Punkte von N_α richtig; denn dort verschwinden die φ_i, die Form $\sum \varphi_{ij}\,\xi_i\,\xi_j$ ist positiv semidefinit und $F'(\alpha) \geq 0$. Wir können daher $\varphi(x) > \alpha$ und folglich $F'(\varphi) > 0$ voraussetzen und statt (10) die Form $Q_\sigma(\xi, \xi)$ mit $\sigma = F''/F'$ betrachten. Nun ist nach der obigen Definition von $F(t)$

$$\sigma = \frac{F''(\varphi)}{F'(\varphi)} \geq G(\varphi) \geq \sigma_0 = -\frac{S_r}{k^2\,S^*_{r-1}},$$

also

$$Q_\sigma(\xi, \xi) \geq Q_{\sigma_0}(\xi, \xi).$$

Es genügt also zu zeigen, daß die letztere Form positiv semidefinit ist. Mit den früheren Bezeichnungen haben wir für diese Form

$$T_\nu^{(\sigma_0)} = S_\nu - \frac{S_r}{S^*_{r-1}}\,S^*_{\nu-1}, \qquad \nu = 1, \ldots, n,$$

also

$$T_\nu^{(\sigma_0)} = 0 \qquad \text{für} \qquad r \leq \nu \leq n$$

auf Grund der Definition von r und wegen III. Die Form Q_{σ_0} hat also höchstens den Rang $r-1$. Da sie nun für $\sum \varphi_i\,\xi_i = 0$ mit P, also mit P^* übereinstimmt und nach II die letztere Form $r-1$ positive Eigenwerte hat, hat sie selbst $r-1$ positive Eigenwerte und ist daher nichtnegativ, wie behauptet.

5. Beschränktheits- und Unbeschränktheitsrichtungen der Niveaumannigfaltigkeiten konvexer Funktionen. Wie schon in der Einleitung erwähnt, haben die Niveaumannigfaltigkeiten einer konvexen Funktion alle dieselben unendlich fernen Punkte. Dies muß folglich auch für die gegebene Funktion $\varphi(x)$ gelten und daher eine Folge der aufgestellten notwendigen Bedingungen sein. Es läßt sich aber einfach direkt und ohne Differenzierbarkeitsvoraussetzungen zeigen. Zur Beleuchtung der Natur der globalen Bedingungen möge der Beweis dafür und für die damit eng zusammenhängende Eigenschaft der Niveaumannigfaltigkeiten, in denselben Richtungen beschränkt zu sein, hier kurz angegeben werden.

Vorbereitend stellen wir einige, unter teilweise engeren Voraussetzungen im wesentlichen bekannte Sätze über unbeschränkte konvexe Mengen mit kurzen Beweisen zusammen (vgl. hierzu [5], [6], [2]).

Es sei zunächst M eine beliebige nicht beschränkte Punktmenge und $u = (u_1, \ldots, u_n) \neq (0, \ldots, 0)$ ein Vektor im E^n. Wir sagen, M sei beschränkt in der Richtung u, oder u sei Beschränktheitsrichtung von M, wenn es eine Hyperebene $\sum u_i\,x_i = c$ mit der positiven Normalenrichtung u gibt, derart daß M in dem durch sie bestimmten abgeschlossenen negativen Halbraum enthalten ist, wenn also

$$\sum u_i\,x_i \leq c \qquad \text{für} \qquad x \in M.$$

Die Hyperebene heiße dann eine Schranke von M; sie kann speziell Stützhyperebene von M sein. Die sämtlichen Beschränktheitsrichtungen u von M machen, etwa vom Nullpunkt aus abgetragen, einen konvexen Kegel $B(M)$ aus, der speziell leer sein kann. Sind nämlich u und v Beschränktheitsrichtungen, so auch $\lambda u+\mu v$ für alle nichtnegativen λ und μ, für welche $\lambda u+\mu v \neq 0$ ist.

Sei nun $x^{(i)}$, $i=1, 2, \ldots$, eine Folge zu M gehöriger Punkte, die sich im Endlichen nicht häufen. Die Folge der Halbstrahlen $\vec{p x}^{(i)}$, die einen willkürlichen festen Punkt p mit den Punkten $x^{(i)}$ verbinden, häuft sich gegen wenigstens einen Halbstrahl. Die Menge aller solchen Häufungsstrahlen ist ein Kegel $A_p(M)$ mit dem Scheitel p. Er gibt die Richtungen an, in denen sich M ins Unendliche erstreckt, und heiße der Asymptotenkegel von M mit dem Scheitel p. Man sieht sehr leicht, daß $A_p(M)$ abgeschlossen ist und daß zwei zu verschiedenen Punkten p gehörige Kegel $A_p(M)$ durch Translation auseinander hervorgehen. Wir können daher kurz von dem Asymptotenkegel von M sprechen.

Zwischen den beiden eingeführten Kegeln besteht die Relation

$$\overline{B(M)} \subseteq B\big(A(M)\big), \tag{11}$$

wo die Überstreichung den Übergang zur abgeschlossenen Hülle bedeutet; denn eine Schranke von M ist, wie unmittelbar aus der Definition von $A_p(M)$ folgt, auch Schranke des Asymptotenkegels mit einem in ihr gelegenen Scheitel, und der Kegel B der Beschränktheitsrichtungen eines Kegels ist abgeschlossen.

Setzen wir nun M als konvex und (der Übersichtlichkeit halber) als abgeschlossen voraus, so ergeben sich wesentlich schärfere Aussagen. Zunächst folgt, daß dann auch $A_p(M)$ konvex ist. Wird ferner p beliebig in M gewählt, so besteht $A_p(M)$ genau aus den von p ausgehenden Halbstrahlen, die ganz zu M gehören. Daß diese Halbstrahlen in $A_p(M)$ enthalten sind, ist nämlich klar. Ist andererseits $\vec{p x}^{(i)}$ eine Folge von Halbstrahlen der oben beschriebenen Art, so gehören die Strecken $p x^{(i)}$ wegen der Konvexität von M ganz zu M. Da nun die Längen dieser Strecken über alle Grenzen wachsen, ist also jeder Punkt eines Halbstrahls, gegen den sich die Halbstrahlen $\vec{p x}^{(i)}$ häufen, Häufungspunkt von Punkten aus M und gehört somit wegen der Abgeschlossenheit von M zu M.

Ferner läßt sich die Inklusion (11) für abgeschlossene konvexe M zu

$$\overline{B(M)} = B\big(A(M)\big) \tag{12}$$

verschärfen. Um dies einzusehen, bemerke man zunächst, daß es genügt, den Fall zu betrachten, wo $A(M)$ keinen vollständigen linearen Unterraum positiver Dimension enthält. Sei nämlich L_o der maximale in $A_o(M)$ enthaltene, durch den Nullpunkt o gehende lineare Unterraum, l seine Dimension und L_o^* der durch o gehende, zu L_o vollständig orthogonale $(n-l)$-dimensionale Unterraum. Der Polarkegel $B\big(A_o(M)\big)$ von $A_o(M)$ ist dann in L_o^* enthalten und hat als Teilmenge dieses Unterraums innere Punkte. Für jeden Punkt p

von M gilt wegen der genannten Eigenschaften der Kegel $A_p(M)$, daß der durch p gehende, zu L_o parallele l-dimensionale Unterraum ganz in M enthalten ist; mit anderen Worten, M ist die Vereinigung von zu L_o parallelen l-dimensionalen Unterräumen. Dies hat zur Folge, daß auch $B(M)$ in L_o^* enthalten ist und daß die Kegel $B(M)$ und $B(A_o(M))$ ungeändert bleiben, wenn man statt E^n den Unterraum L_o^* betrachtet und M und somit $A_o(M)$ durch ihre Durchschnitte mit L_o^* ersetzt. Dies zeigt, daß wir ohne Beschränkung der Allgemeinheit zu den obigen Voraussetzungen hinzufügen können, daß $A(M)$ keinen linearen Unterraum positiver Dimension vollständig enthält. Es gibt dann Stützhyperebenen des Kegels $A_o(M)$, die mit ihm nur den Punkt o gemeinsam haben. Die positiven Normalen solcher Stützhyperebenen sind genau die inneren Halbstrahlen des Polarkegels $B(A_o(M))$. Zum Beweis von (12) genügt es nun wegen (11) zu zeigen, daß ein nicht zu $B(M)$ gehöriger Halbstrahl nicht innerer Strahl von $B(A_o(M))$ sein kann. Dies ist aber evident. Ist nämlich u ein nicht zu $B(M)$ gehöriger Vektor, so enthält der Halbraum $\sum u_i x_i > 0$ eine ins Unendliche strebende Folge von Punkten $x^{(i)}$ aus M. Die Häufungsstrahlen der Folge von Halbstrahlen $o\vec{x}^{(i)}$ sind folglich im Halbraum $\sum u_i x_i \geq 0$ enthalten, und $A_o(M)$ kann somit (vom Scheitel o abgesehen) nicht vollständig dem offenen Halbraum $\sum u_i x_i < 0$ angehören.

Wir wenden uns nun zur Anwendung dieser Resultate auf die Niveaumannigfaltigkeiten konvexer Funktionen. Sei also $z = f(x)$ eine im offenen konvexen Bereich C definierte konvexe Funktion. Mit $[C, f]$ werde die abgeschlossene und konvexe Menge derjenigen Punkte (x, z) des E^{n+1} bezeichnet, für welche $x \in C$ und $z \geq f(x)$. Es sei M_{z_0} der Durchschnitt von $[C, f]$ mit der Hyperebene $z = z_0$. Die orthogonale Projektion von M_{z_0} auf die Hyperebene $z = 0$ ist die Niveaumenge N_{z_0} von f, d.h. die Menge der Punkte x mit $f(x) \leq z_0$. Der Asymptotenkegel $A_{p_0}[C, f]$, wo $p_0 = (x^{(0)}, z_0)$ ein Punkt von $[C, f]$ ist, wird von der Hyperebene $z = z_0$ im Asymptotenkegel $A_{p_0}(M_{z_0})$ von M_{z_0} geschnitten. Ist $p_1 = (x^{(1)}, z_1)$ ein zweiter Punkt von $[C, f]$, so geht $A_{p_1}[C, f]$ aus $A_{p_0}[C, f]$ durch eine Translation hervor. Dieselbe Translation führt aber die Hyperebene $z = z_0$ in $z = z_1$ über, und folglich ist $A(M_{z_0}) = A(M_{z_1})$. Projiziert man auf $z = 0$, so erhält man also die Behauptung:

Die Niveaumengen einer konvexen Funktion sind entweder alle beschränkt oder alle unbeschränkt mit demselben Asymptotenkegel.

Aus (12) kann man nun weiter schließen, daß die Kegel $B(N_z)$ der Beschränktheitsrichtungen für alle Niveaumengen N_z dieselbe abgeschlossene Hülle haben. Es läßt sich aber noch etwas mehr aussagen:

Hat eine konvexe Funktion $f(x)$ die untere Grenze a $(\geq -\infty)$, so sind alle Niveaumengen N_z für $z > a$ in denselben Richtungen beschränkt. Ist $B = B(N_z)$ ihr gemeinsamer Kegel der Beschränktheitsrichtungen, so gilt, falls $f(x)$ die untere Grenze a erreicht, $B \subseteq B(N_a) \subseteq \overline{B}$.

Die zweite Behauptung folgt unmittelbar aus $N_a \subseteq N_z$ für $z > a$ und dem soeben über die Kegel $B(N_z)$ Gesagten. Die Richtigkeit der ersten Behauptung kann man so einsehen. Es sei N_{z_0}, $z_0 > a$, in der Richtung u beschränkt.

Ist $\sum u_i x_i = c$ eine Schranke, so ist der $(n-1)$-dimensionale Unterraum L^{n-1} des E^{n+1} mit den Gleichungen $\sum u_i x_i = c$, $z = z_0$ eine Schranke der Menge M_{z_0} als Teilmenge der Hyperebene $z = z_0$. Dieser Unterraum L^{n-1} enthält also keine inneren Punkte der Menge $[C, f]$. Daher läßt sich durch ihn eine Schranke von $[C, f]$ legen. Durch Projektion auf eine zu L^{n-1} orthogonale (zweidimensionale) Ebene des E^{n+1} geht nämlich $[C, f]$ in einen (unbeschränkten) konvexen Bereich und L^{n-1} in einen äußeren oder Randpunkt dieses Bereichs über. Durch diesen Punkt läßt sich folglich eine Gerade legen, die Schranke des Bereichs ist und die daher zusammen mit L^{n-1} eine Schranke S von $[C, f]$ aufspannt. Nun muß S von der Hyperebene $z = z_0$ verschieden sein, da $[C, f]$ sowohl Punkte mit $z > z_0$ als auch Punkte mit $z < z_0$ enthält. Daher wird S von den Hyperebenen $z = \text{konst.}$ in zu L^{n-1} parallelen Schranken der Mengen M_z geschnitten. Projektion auf $z = 0$ ergibt also, daß jede Beschränktheitsrichtung von N_{z_0} auch Beschränktheitsrichtung aller übrigen Mengen N_z ist. Damit ist die Behauptung bewiesen.

Daß im Falle der Existenz eines Minimums von $f(x)$ der Kegel $B(N_a)$ der Menge N_a der Punkte, in denen das Minimum angenommen wird, umfassender sein kann als der Kegel B der Beschränktheitsrichtungen der übrigen Niveaumengen, zeigt die konvexe Funktion $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2} - x_1$. Hier ist $a = 0$, und N_0 ist die positive x_1-Achse, also beschränkt in den Richtungen (u_1, u_2) mit $u_1 \leq 0$, während die anderen Mengen N_z von Parabeln mit der x_1-Achse als Achse begrenzt werden, also nur in Richtungen mit $u_1 < 0$ beschränkt sind.

Beispiele von Funktionen $\varphi(x_1, x_2)$, die die Bedingungen I bis III erfüllen, deren Niveaukurven aber nicht die soeben bewiesenen Eigenschaften der Niveaumannigfaltigkeiten konvexer Funktionen besitzen, sind ebenfalls leicht anzugeben. Die Niveaukurven der Funktion $\varphi(x_1, x_2) = (1 + x_2^2)/x_1^2$, $x_1 > 0$, sind die in $x_1 > 0$ gelegenen Zweige der konzentrischen Hyperbeln $t x_1^2 - x_2^2 = 1$, $t > 0$, mit von t abhängigem Asymptotenwinkel. Bei der Funktion $\varphi(x_1, x_2) = [(x_1 - 1)^2 + x_2^2]/x_1^2$, $x_1 > 0$, sind es die Kegelschnitte mit dem Brennpunkt $(1, 0)$ und der Leitlinie $x_1 = 0$. Die Schar enthält also hier sowohl beschränkte als auch unbeschränkte Kurven. In beiden Fällen ist die Bedingung IV nicht erfüllt.

Literatur

[1] COURANT, R., u. D. HILBERT: Methoden der mathematischen Physik, 2. Aufl., Bd. I. Berlin 1931. — [2] FENCHEL, W.: Convex cones, sets, and functions. (From notes by D. W. BLACKETT of lectures at Princeton University.) Princeton, N. J. 1953. — [3] FINETTI, B. DE: Sulle stratificazioni convesse. Ann. di Matematica (4) **30**, 173—183 (1949). — [4] KOWALEWSKI, G.: Einführung in die Determinantentheorie. Leipzig 1909. — [5] STEINITZ, E.: Bedingt konvergente Reihen und konvexe Systeme, II. J. reine u. angew. Math. **144**, 1—40 (1914). — [6] STOKER, J. J.: Unbounded convex point sets. Amer. J. Math. **62**, 165—179 (1940).

Kopenhagen, Dänische Technische Hochschule

(Eingegangen am 23. September 1955)

KOCHENDÖRFFER, R.
Math. Zeitschr. Bd. 63, S. 507—513 (1956)

Über den Multiplikator einer Gruppe

Zum Gedenken an den 80. Geburtstag meines verehrten Lehrers I. SCHUR

Von

RUDOLF KOCHENDÖRFFER

In der von I. SCHUR[1]) begründeten Theorie der Darstellungen endlicher Gruppen durch gebrochene lineare Substitutionen wird jeder endlichen Gruppe $\mathfrak{H}$ ihr *Multiplikator* $\mathfrak{M} = \mathfrak{M}(\mathfrak{H})$ zugeordnet. Dieser spielt eine entscheidende Rolle bei der Aufgabe, eine Übersicht über alle Darstellungen von $\mathfrak{H}$ zu gewinnen. Der Multiplikator ist eine abelsche Gruppe, in deren Ordnung höchstens diejenigen Primzahlen aufgehen, die die Ordnung von $\mathfrak{H}$ teilen. Die vollständige Berechnung des Multiplikators einer gegebenen Gruppe ist in den meisten Fällen eine schwierige Aufgabe, so daß bereits Teilaussagen über seine Struktur von Interesse sein können. Eine solche Teilaussage ist die folgende:

Die p-Sylowgruppe des Multiplikators $\mathfrak{M}$ von $\mathfrak{H}$ ist isomorph zu einer Untergruppe des Multiplikators einer p-Sylowgruppe von $\mathfrak{H}$[2]).

Welche Untergruppe ihres Multiplikators eine Sylowgruppe in diesem Sinne zum Multiplikator der ganzen Gruppe beisteuert, hängt von ihrer Einbettung in die ganze Gruppe ab. Eine Bedingung dafür, daß eine abelsche Sylowgruppe überhaupt keinen Beitrag zum Multiplikator liefert, findet sich in S II[3]).

Im folgenden soll der oben formulierte Satz aufs neue bewiesen werden, und zwar mit einem Verfahren, welches eine Aussage darüber gestattet, wann eine Sylowgruppe ihren vollen Multiplikator zum Multiplikator der ganzen Gruppe beisteuert. Es ergibt sich folgender Satz:

Es sei $\mathfrak{S}$ eine p-Sylowgruppe von $\mathfrak{H}$ und

$$\mathfrak{H} = \sum_{V \in \mathfrak{V}} \mathfrak{S} V,$$

wobei V ein Rechtsvertretersystem $\mathfrak{V}$ von $\mathfrak{H}$ nach $\mathfrak{S}$ durchläuft. Besitzt der Komplex $\mathfrak{V}$ die Eigenschaft

$$P^{-1} \mathfrak{V} P = \mathfrak{V} \qquad \text{für jedes } P \text{ aus } \mathfrak{S},$$

so ist die p-Sylowgruppe des Multiplikators von $\mathfrak{H}$ isomorph zum Multiplikator von $\mathfrak{S}$.

[1]) I. SCHUR, Über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **127**, 20—50 (1904). — Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **132**, 85—137 (1907). Im folgenden zitiert mit SI und SII.

[2]) Vgl. SI, Satz IX.

[3]) SII, Seite 114.

Die Voraussetzung dieses Satzes ist insbesondere erfüllt, wenn $\mathfrak{S}$ direkter Faktor ist. In diesem Spezialfall folgt unsere Aussage aus einem Satz von SCHUR[4]).

Als Koeffizientenbereich für alle Substitutionen werde ein algebraisch abgeschlossener Körper zugrunde gelegt, dessen Charakteristik kein Teiler der Ordnung von $\mathfrak{H}$ ist. Ist

$$(A) = (a_{ik}) \qquad (i, k = 1, \ldots, n)$$

die Koeffizientenmatrix der dem Element A aus $\mathfrak{H}$ zugeordneten gebrochenen linearen Substitution

$$L_A: \qquad x_\nu = \frac{a_{\nu 1} y_1 + \cdots + a_{\nu, n-1} y_{n-1} + a_{\nu n}}{a_{n1} y_1 + \cdots + a_{n, n-1} y_{n-1} + a_{nn}} \qquad (\nu = 1, \ldots, n-1),$$

so gilt für je zwei Elemente A, B aus $\mathfrak{H}$

$$(A)\,(B) = r_{A,B}\,(AB) \tag{1}$$

mit einem gewissen Faktorensystem $\{r_{A,B}\}$ aus nicht verschwindenden Zahlen. Infolge des assoziativen Gesetzes in $\mathfrak{H}$ genügen die Zahlen $r_{A,B}$ den Relationen

$$r_{A,BC}\, r_{B,C} = r_{AB,C}\, r_{A,B}. \tag{2}$$

Umgekehrt gehören zu jedem Faktorensystem $\{r_{A,B}\}$, welches die Bedingungen (2) erfüllt, Matrizen $(A), (B), \ldots$, die den Gln. (1) genügen und somit eine Darstellung von $\mathfrak{H}$ durch gebrochene lineare Substitutionen bilden.

Ersetzt man die Matrizen (A) durch proportionale Matrizen $c_A \cdot (A)$ mit beliebigen Zahlfaktoren $c_A \neq 0$, so erhält man die gleichen gebrochenen linearen Substitutionen, während das Faktorensystem $\{r_{A,B}\}$ in das System

$$\{r'_{A,B}\} = \left\{\frac{c_A c_B}{c_{AB}} \cdot r_{A,B}\right\} \tag{3}$$

übergeht. Zwei Faktorensysteme $\{r_{A,B}\}$ und $\{r'_{A,B}\}$, die durch (3) miteinander zusammenhängen, heißen assoziiert, in Zeichen:

$$\{r_{A,B}\} \sim \{r'_{A,B}\}.$$

Ein Faktorensystem, welches zum Einssystem $\{1\}$ assoziiert ist, wird als ein zerfallendes Faktorensystem bezeichnet.

Hat man zwei Faktorensysteme $\{r_{A,B}\}$ und $\{s_{A,B}\}$, die (2) erfüllen, so genügen offenbar auch das System $\{r^{-1}_{A,B}\}$ sowie das Produktsystem

$$\{r_{A,B}\}\,\{s_{A,B}\} = \{r_{A,B} \cdot s_{A,B}\}$$

den Bedingungen (2) und gehören infolgedessen ebenfalls zu Darstellungen von $\mathfrak{H}$ durch gebrochene lineare Substitutionen. Da die Assoziiertheit eine mit dieser Multiplikation der Faktorensysteme verträgliche Äquivalenzrelation darstellt, bilden die Klassen assoziierter Faktorensysteme eine abelsche Gruppe

[4]) SII, Satz VII.

$\mathfrak{M}$, die als der *Multiplikator* von $\mathfrak{H}$ bezeichnet wird. Es zeigt sich, daß $\mathfrak{M}$ eine endliche Ordnung besitzt, in der nur solche Primzahlen aufgehen, die die Ordnung von $\mathfrak{H}$ teilen.

Die Grundlage für das Folgende bildet eine Schlußweise von W. GASCHÜTZ[5]), die sich bei der Untersuchung von Gruppenerweiterungen bewährt hat.

Für die Zerlegung von $\mathfrak{H}$ in Rechtsnebenklassen nach $\mathfrak{S}$ legen wir ein festes Vertretersystem $\mathfrak{V}$ zugrunde und haben dann

$$\mathfrak{H} = \sum_{V \in \mathfrak{V}} \mathfrak{S} V. \tag{4}$$

Sei $\{r'_{A,B}\}$ irgendein zur Gruppe $\mathfrak{H}$ gehöriges Faktorensystem. Das Teilsystem $\{r'_{P,Q}\}$ für die Elementepaare $P, Q \in \mathfrak{S}$ bildet dann ein Faktorensystem für die Sylowgruppe $\mathfrak{S}$ von $\mathfrak{H}$. Wir wollen nun zunächst das Faktorensystem $\{r'_{A,B}\}$ durch ein assoziiertes System

$$\{r_{A,B}\} = \left\{ \frac{c_{AB}}{c_A c_B} \cdot r'_{A,B} \right\}$$

ersetzen, und zwar durch folgende Wahl der Zahlen c_A:

$$c_P = 1 \quad \text{für} \quad P \in \mathfrak{S},$$

$$c_V \neq 0, \text{ beliebig für } V \in \mathfrak{V},$$

$$c_{PV} = c_P c_V r'^{-1}_{P,V}.$$

Dadurch wird erreicht, daß das Teilfaktorensystem $\{r'_{P,Q}\}$ mit $P, Q \in \mathfrak{S}$ ungeändert bleibt, während für $P \in \mathfrak{S}$ und $V \in \mathfrak{V}$

$$r_{P,V} = \frac{c_{PV}}{c_P c_V} \cdot r'_{P,V} = \frac{c_P c_V r'^{-1}_{P,V}}{c_P c_V} r'_{P,V} = 1 \tag{5}$$

wird.

Um die jetzt vorzunehmenden Umformungen übersichtlich zu gestalten, empfiehlt es sich, folgende Bezeichnungsweise aus [5]) zu übernehmen:

Ein beliebiges Element H aus $\mathfrak{H}$ besitzt gemäß (4) eine eindeutige Zerlegung

$$H = PV \quad \text{mit} \quad P \in \mathfrak{S},\ V \in \mathfrak{V}.$$

Wir setzen

$$P = \underline{H},\ V = \overline{H}, \quad \text{also} \quad H = \underline{H}\overline{H}.$$

Die Relationen (2) benutzen wir im Spezialfall $A = P \in \mathfrak{S}$, $B = \underline{H}$, $C = \overline{H}$ und erhalten

$$r_{P,\underline{H}\overline{H}} r_{\underline{H},\overline{H}} = r_{P\underline{H},\overline{H}} r_{P,\underline{H}},$$

also wegen (5)

$$r_{P,H} = r_{P,\underline{H}\overline{H}} = r_{P,\underline{H}} \qquad (P \in \mathfrak{S},\ H \in \mathfrak{H}). \tag{6}$$

Ein zweites Mal benutzen wir (2) mit $A = \underline{H}$, $B = \overline{H}$, $C \in \mathfrak{H}$:

$$r_{\underline{H},\overline{H}C} r_{\overline{H},C} = r_{\underline{H}\overline{H},C} r_{\underline{H},\overline{H}} = r_{H,C} r_{\underline{H},\overline{H}}.$$

[5]) GASCHÜTZ, W.: Zur Erweiterungstheorie der endlichen Gruppen. J. reine u. angew. Math. **190**, 93—107 (1952).

Hieraus folgt unter Berücksichtigung von (5)

$$r_{\underline{H},\overline{H}C}\, r_{\overline{H},C} = r_{H,C}$$

und weiter, indem man (6) auf den ersten Faktor der linken Seite anwendet,

$$r_{\underline{H},\underline{\overline{H}C}}\, r_{\overline{H},C} = r_{H,C} \qquad (C, H \in \mathfrak{H}). \tag{7}$$

Die Relation (2) läßt sich in der Form

$$r_{B,C} = r^{-1}_{A,\overline{B}C}\, r_{A,B}\, r_{\underline{A}B,C}$$

schreiben. Den letzten Faktor rechts formen wir mittels (7) um und erhalten

$$r_{B,C} = r^{-1}_{A,\overline{B}C}\, r_{A,B}\, r_{\underline{A}B,C}\, r_{\underline{A}B,\underline{\overline{A}B}\,C}.$$

Hierin lasse man A die $k = [\mathfrak{H}:\mathfrak{S}]$ Elemente des Vertretersystems $\mathfrak{V}$ durchlaufen und multipliziere alle diese Gleichungen miteinander. Das ergibt

$$r^k_{B,C} = \prod_{V\in\mathfrak{V}} r^{-1}_{V,\overline{B}C} \cdot \prod_{V\in\mathfrak{V}} r_{V,B} \cdot \prod_{V\in\mathfrak{V}} r_{V,C} \cdot \prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{V}BC}},$$

wobei man für die dritten Faktoren auf der rechten Seite nur zu beachten hat, daß $\overline{VB}$ bei festem B zugleich mit V alle Elemente aus $\mathfrak{V}$ durchläuft. Setzt man zur Abkürzung

$$\prod_{V\in\mathfrak{V}} r_{V,H} = a_H,$$

so erhält man

$$r^k_{B,C} = \frac{a_B a_C}{a_{BC}} \cdot \prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{V}BC}}. \tag{8}$$

Das Zahlensystem

$$\{s_{B,C}\} = \{\prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{V}BC}}\} \tag{9}$$

hat sich mithin als ein zu $\{r^k_{B,C}\}$ assoziiertes Faktorensystem erwiesen.

Jedem zur vollen Gruppe $\mathfrak{H}$ gehörigen Faktorensystem

$$\{r_{A,B}\} \qquad (A, B \in \mathfrak{H})$$

läßt sich ein zur Untergruppe $\mathfrak{S}$ gehöriges Faktorensystem zuordnen, nämlich einfach das Teilsystem

$$\{r_{P,Q}\} \qquad (P, Q \in \mathfrak{S}).$$

Assoziierten $\mathfrak{H}$-Faktorensystemen entsprechen bei dieser Zuordnung selbstverständlich assoziierte $\mathfrak{S}$-Faktorensysteme. Das Produkt zweier $\mathfrak{H}$-Faktorensysteme wird abgebildet auf das Produkt der zu $\mathfrak{S}$ gehörigen Teilsysteme. Man erhält mithin eine homomorphe Abbildung des Multiplikators $\mathfrak{M}$ von $\mathfrak{H}$ auf eine Untergruppe $\mathfrak{U}$ des Multiplikators $\mathfrak{N}$ von $\mathfrak{S}$. Wir bestimmen den Kern dieses Homomorphismus. Da die Ordnung von $\mathfrak{N}$, wie oben bemerkt wurde, eine Potenz von p ist, gehören sicher alle diejenigen Elemente von $\mathfrak{M}$ zum Kern, deren Ordnungen zu p teilerfremd sind. Folglich brauchen nur noch die Elemente der p-Sylowgruppe $\mathfrak{P}$ von $\mathfrak{M}$ weiter untersucht zu werden.

Es wird sich zeigen, daß kein von Eins verschiedenes Element aus $\mathfrak{P}$ zum Kern des betrachteten Homomorphismus gehört oder, mit anderen Worten, daß für die Faktorensysteme aus den $\mathfrak{P}$ angehörenden Klassen aus dem Zerfallen des zu $\mathfrak{S}$ gehörenden Teilsystems bereits das Zerfallen des gesamten Systems folgt.

Um den Nachweis hierfür zu erbringen, kann (8) dienen, weil in dem Produkt auf der rechten Seite nur Zahlen aus dem zu $\mathfrak{S}$ gehörigen Teilsystem auftreten. Sei also $\{r_{A,B}\}$ ein $\mathfrak{H}$-Faktorensystem aus einer der zu $\mathfrak{P}$ gehörenden Klassen, für welches das der Untergruppe $\mathfrak{S}$ entsprechende Teilsystem zerfällt, d.h.

$$r_{P,Q} = \frac{d_P d_Q}{d_{PQ}} \qquad (P, Q \in \mathfrak{S}).$$

Für das Zahlensystem (9) erhält man dann

$$\prod_{V \in \mathfrak{V}} r_{\underline{VB}, \overline{VB}C} = \frac{\prod_{V \in \mathfrak{V}} d_{\underline{VB}} \cdot \prod_{V \in \mathfrak{V}} d_{\overline{VB}C}}{\prod_{V \in \mathfrak{V}} d_{\underline{VB}\,\overline{VB}C}}.$$

Mit der Abkürzung

$$\prod_{V \in \mathfrak{V}} d_{\underline{VB}} = e_B$$

wird

$$\prod_{V \in \mathfrak{V}} d_{\overline{VB}C} = \prod_{V \in \mathfrak{V}} d_{\underline{VC}} = e_C$$

und

$$\prod_{V \in \mathfrak{V}} d_{\underline{VB}\,\overline{VB}C} = \prod_{V \in \mathfrak{V}} d_{\underline{VBC}} = e_{BC}.$$

Bei der letzten Umformung ist von der Gleichung

$$\underline{VB}\,\overline{VB}C = \underline{VBC} \tag{10}$$

Gebrauch gemacht worden, die man leicht folgendermaßen bestätigt: Es werde gesetzt:

$$B = S_1 V_1, \quad C = S_2 V_2 \qquad (S_1, S_2 \in \mathfrak{S};\ V_1, V_2 \in \mathfrak{V}).$$

Dann wird

$$VB = V S_1 V_1 = S_1' V' V_1 = S_1' S W$$

mit

$$V S_1 = S_1' V' \quad \text{und} \quad V' V_1 = S W \qquad (S_1',\ S \in \mathfrak{S};\ V', W \in \mathfrak{V}).$$

Folglich ist

$$\underline{VB} = S_1' S, \quad \overline{VB} = W.$$

Ferner erhält man

$$\underline{\overline{VB}C} = \underline{W S_2 V_2} = \underline{S_2' W' V_2} = \underline{S_2' \overline{S}\,\overline{W}} = S_2' \overline{S},$$

wenn

$$W S_2 = S_2' W' \quad \text{und} \quad W' V_2 = \overline{S}\,\overline{W} \qquad (S_2', \overline{S} \in \mathfrak{S};\ W', \overline{W} \in \mathfrak{V})$$

gesetzt wird. Zusammen ist also

$$\underline{VB}\,\overline{VB}C = S_1' S S_2' \overline{S}.$$

Andererseits ist

$$VBC = \underline{VS_1V_1S_2V_2} = \underline{S_1'V'V_1S_2V_2} = \underline{S_1'SWS_2V_2}$$
$$= \underline{S_1'SS_2'W'V_2} = \underline{S_1'SS_2'\bar{S}\bar{W}} = S_1'SS_2'\bar{S},$$

also in der Tat

$$\underline{VB}\,\underline{\overline{VB}C} = \underline{VBC}.$$

Im ganzen erhalten wir daher für die Zahlen des Systems (9)

$$\prod_{V\in\mathfrak{B}} r_{\underline{VB},\underline{\overline{VB}C}} = \frac{e_B e_C}{e_{BC}}.$$

Daher besagt (8), daß das System $\{r_{A,B}\}^k$ zerfällt. Da der Exponent $k = [\mathfrak{H}:\mathfrak{S}]$ zur Ordnung der das Faktorensystem $\{r_{A,B}\}$ enthaltenden Klasse aus $\mathfrak{P}$ teilerfremd ist, zerfällt $\{r_{A,B}\}$ selbst.

Mithin hat sich ergeben, daß $\mathfrak{P}$ isomorph einer Untergruppe $\mathfrak{U}$ des Multiplikators von $\mathfrak{S}$ ist. Damit ist der erste der oben genannten Sätze bewiesen.

Die Untergruppe $\mathfrak{U}$ des Multiplikators von $\mathfrak{S}$ kann folgendermaßen gekennzeichnet werden: $\mathfrak{U}$ besteht genau aus allen denjenigen Klassen von $\mathfrak{S}$-Faktorensystemen, die als Teilsysteme von $\mathfrak{H}$-Faktorensystemen aufgefaßt werden können oder, wie wir auch sagen wollen, die sich zu $\mathfrak{H}$-Faktorensystemen fortsetzen lassen. Formel (8) lehrt, daß die Fortsetzung eines $\mathfrak{S}$-Faktorensystems zu einem $\mathfrak{H}$-Faktorensystem — falls überhaupt — bis auf assoziierte Systeme nur auf eine einzige Weise möglich ist. Denn sind $\{q_{A,B}\}$ und $\{q'_{A,B}\}$ zwei Fortsetzungen ein und desselben $\mathfrak{S}$-Faktorensystems, so zerfällt von dem System $\{q'_{A,B}\cdot q_{A,B}^{-1}\}$ der zu $\mathfrak{S}$ gehörige Teil, woraus man wie oben schließt, daß das gesamte System zerfällt. Wie die Fortsetzung eines $\mathfrak{S}$-Faktorensystems aussehen muß, falls sie überhaupt existiert, ist ebenfalls aus (8) zu ersehen. Man bestimme eine Lösung l der Kongruenz

$$kl \equiv 1 \pmod{[\mathfrak{P}:1]},$$

was stets möglich ist, da $k = [\mathfrak{H}:\mathfrak{S}]$ zu der p-Potenz $[\mathfrak{P}:1]$ teilerfremd ist. Dann folgt aus (8), daß das Faktorensystem $\{r_{B,C}\}$ zu dem System

$$\{t_{B,C}\} = \{s_{B,C}\}^l = \left\{\prod_{V\in\mathfrak{B}} r_{\underline{VB},\underline{\overline{VB}C}}\right\}^l$$

assoziiert ist. Falls also das $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ überhaupt eine Fortsetzung besitzt, so muß diese zu $\{t_{B,C}\}$ assoziiert sein.

Das Zahlensystem $\{t_{B,C}\}$ läßt sich nun für jedes beliebige $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ bilden. Man bestätigt ohne Mühe, daß $\{t_{B,C}\}$ stets ein $\mathfrak{H}$-Faktorensystem darstellt, indem man die Richtigkeit der Gleichungen

$$t_{A,BC}\,t_{B,C} = t_{AB,C}\,t_{A,B}$$

nachprüft. Dabei hat man neben (10) von der Beziehung

$$\overline{VA}\,\overline{B} = \overline{VAB}$$

Gebrauch zu machen, die man ähnlich wie (10) leicht beweist. Hat man für ein beliebiges $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ das System $\{t_{B,C}\}$ gebildet, so bleibt zu untersuchen, ob dadurch tatsächlich eine Fortsetzung von $\{r_{P,Q}\}$ gewonnen ist, d.h. ob die beiden $\mathfrak{S}$-Faktorensysteme $\{r_{P,Q}\}$ und $\{t_{P,Q}\}$ assoziiert sind. Falls jedes $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ zu dem entsprechenden System $\{t_{P,Q}\}$ assoziiert ist, kann jedes $\mathfrak{S}$-Faktorensystem zu einem $\mathfrak{H}$-Faktorensystem fortgesetzt werden. Die Beziehung

$$\{r_{P,Q}\} \sim \{t_{P,Q}\}$$

ist gleichbedeutend mit

$$\{r_{P,Q}\}^k \sim \{t_{P,Q}\}^k$$

und wegen

$$\{t_{P,Q}\}^k \sim \{s_{P,Q}\} = \left\{ \prod_{V \in \mathfrak{V}} r_{\underline{VP},\overline{VP}Q} \right\}$$

auch gleichbedeutend mit dem Zerfallen des Systems

$$\left\{ \prod_{V \in \mathfrak{V}} \frac{r_{\underline{VP},\overline{VP}Q}}{r_{P,Q}} \right\}.$$

Dieses System zerfällt sicher, wenn jeder einzelne Faktor des Produktes gleich 1 ist, d.h.

$$r_{\underline{VP},\overline{VP}Q} = r_{P,Q} \qquad (P, Q \in \mathfrak{S};\ V \in \mathfrak{V}).$$

Das tritt jedenfalls ein, wenn für jedes Paar $P \in \mathfrak{S}$, $V \in \mathfrak{V}$

$$\underline{VP} = P$$

wird oder

$$VP = PP^{-1}VP = PV'$$

mit $V' = P^{-1}VP \in \mathfrak{V}$. Das bedeutet, daß sich ein Vertretersystem $\mathfrak{V}$ wählen läßt, welches die Eigenschaft

$$P^{-1}\mathfrak{V}P = \mathfrak{V} \quad \text{für jedes } P \in \mathfrak{S}$$

besitzt. Damit ist der zweite der oben angegebenen Sätze bewiesen.

Rostock, Mathematisches Institut der Universität

(Eingegangen am 9. August 1955)

GEIRINGER, H.
Math. Zeitschr. Bd. 63, S. 514—524 (1956)

Grenzlinien der Hodographentransformation

ISSAI SCHUR zum Gedächtnis

Von

HILDA GEIRINGER

In der Theorie der ebenen Hodographentransformation der Aerodynamik spielen gewisse Singularitäten dieser Transformation, die dem Verschwinden Jacobischer Funktionaldeterminanten entsprechen, eine wichtige Rolle. Es ist verwunderlich, daß die Diskussion solcher Singularitäten nicht von einem allgemeinen Standpunkt aus in Lehrbüchern der Differential- und Integralrechnung Platz findet. Diese beschränken sich im allgemeinen auf die Bemerkung, daß die Funktionaldeterminante der Abbildung als von Null verschieden vorausgesetzt wird; es wird aber kaum darauf eingegangen, was im Falle ihres Verschwindens geschieht. Vorbereitet durch Arbeiten von RINGLEB [*1*] und TOLLMIEN [*2*] finden sich Studien dieser Richtung, mit speziellem Bezug auf das aerodynamische Problem, unter anderem in einer inhaltsreichen Arbeit von CRAGGS [*3*], in dem bekannten Lehrbuch von COURANT und FRIEDRICHS [*4*] und in dem von HOWARTH [*5*] herausgegebenen Sammelwerk.

Die folgende Darstellung unterscheidet sich von den genannten teils in der Form, teils im Inhalt. Unsere sehr einfache Methode scheint den Methoden der unter [*3*], [*4*], [*5*] genannten Darstellungen überlegen. Auch inhaltlich gehen wir über die in diesen Arbeiten gegebenen Resultate (besonders über die in [*4*] gegebenen) in mancher Hinsicht hinaus. Die hier benutzte Methode wurde in einer mir erst nachträglich bekannt gewordenen Arbeit von R. E. MEYER [*6*] angewendet, die sich aber nicht mit der Darstellung der grundlegenden Eigenschaften der Singularitäten, sondern mit verfeinerten Fragestellungen vom Standpunkt der Riemannschen Geometrie beschäftigt. Hingegen ist die fragliche Methode schon 1941 von J. MANDEL [*7*] angewendet worden, allerdings nicht auf die Hodographentransformation der Aerodynamik, sondern auf das Spannungsproblem der ebenen Plastizitätstheorie. Dieses hat mit jener mathematisch entscheidende Ähnlichkeiten [*8*], [*9*]. Die vorliegende Darstellung ist von der MANDELS methodisch beeinflußt.

Um Raum zu sparen, geben wir unsere Resultate gleich in aerodynamischer Form, ohne erst auf die allgemeine Abbildungsaufgabe einzugehen.

1. Aerodynamische Grundgleichungen. Wir betrachten das ebene stationäre Problem eines kompressiblen, wirbelfreien, reibungsfreien Gases. Dieses wird als „elastisch" angenommen, d. h. es existiert eine universale, eineindeutige Beziehung zwischen Druck p und Dichte ϱ. Seien x, y die cartesischen Koordinaten der Strömungsebene, u, v die x- und y-Komponenten des Geschwindigkeitsvektors $\bar{q}$, dessen Polarkoordinaten q und ϑ sind. Die

u, v-Ebene heißt Hodographenebene oder Hodograph. Bezeichnet $\frac{\partial}{\partial s}$ und $\frac{\partial}{\partial n}$ Differentiation in Richtung von $\bar{q}$ und normal dazu, so ist

$$(1) \qquad \frac{\partial}{\partial s} = \cos\vartheta \frac{\partial}{\partial x} + \sin\vartheta \frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial n} = -\sin\vartheta \frac{\partial}{\partial x} + \cos\vartheta \frac{\partial}{\partial y},$$

und es gelten die Gleichungen

$$(2) \qquad \frac{\partial q}{\partial n} = q \frac{\partial \vartheta}{\partial s}, \qquad (M^2 - 1) \frac{\partial q}{\partial s} = q \frac{\partial \vartheta}{\partial n},$$

die Wirbelfreiheit und Kontinuität ausdrücken. Dabei ist

$$(3) \qquad M^2 = q^2 \frac{d\varrho}{dp}$$

und wir setzen $\frac{dp}{d\varrho} = a^2$; ferner ist q^2 mit p und ϱ durch die Bernoullische Gleichung

$$(4) \qquad \frac{q^2}{2} + \int^p \frac{dp}{\varrho} = \text{constant}$$

verbunden. Mittels (3) und (4) kann M^2 als Funktion von q^2 ausgedrückt werden. Wird die p, ϱ-Beziehung in der adiabatischen Form $p\varrho^{-\gamma} = \text{constant}$ angenommen, so folgt

$$(3') \qquad M^2 = \frac{q^2}{a^2} = \frac{q^2}{a_0^2 - \frac{\gamma - 1}{2} q^2},$$

wo a_0 den Wert der Schallgeschwindigkeit a für $q = 0$ („Stagnationswert" von a) bezeichnet. M^2 wächst monoton mit q^2 von $M = 0$ für $q = 0$, zu $M = \infty$ für die Maximalgeschwindigkeit q_m, für welche p, ϱ und a verschwinden; für $M = 1$ wird die Schallgeschwindigkeit oder „kritische" Geschwindigkeit $q = a$ angenommen. Ist M^2 durch q^2 ausgedrückt, so hat man in (2) zwei Gleichungen für q und ϑ, und p und ϱ folgen sodann mittels (4) und der p, ϱ-Beziehung.

Für $M > 1$ (Überschallgeschwindigkeit) setzt man

$$(5) \qquad M = \frac{1}{\sin\alpha}, \qquad M^2 - 1 = \operatorname{ctg}^2\alpha, \qquad 0 \leq \alpha \leq 90°$$

und es ergibt sich, daß

$$(6) \qquad \varphi^+ = \vartheta + \alpha, \qquad \varphi^- = \vartheta - \alpha$$

die *charakteristischen Richtungen* von (2) sind. Die überall den Winkel φ^- mit der x-Achse einschließenden Kurven werden als C^--Charakteristiken oder als ξ-Kurven, die andern als C^+-Charakteristiken oder η-Kurven bezeichnet. Setzt man

$$(7) \qquad \frac{dQ}{dq} = \frac{1}{q \operatorname{tang}\alpha}, \qquad Q = \int^q \frac{dq}{q \operatorname{tang}\alpha},$$

so hat $Q + \vartheta$ entlang einer C^-, und $Q - \vartheta$ entlang einer C^+ konstanten Wert *(Kompatibilitätsbedingungen)*. Wir können daher

$$(8) \qquad Q + \vartheta = 2\eta, \qquad Q - \vartheta = 2\xi,$$

$$(9) \qquad Q = \eta + \xi, \qquad \vartheta = \eta - \xi$$

setzen.

Die Gln. (2) sind *nichtlinear*, da die Koeffizienten von der abhängigen Variabeln q abhängen. Da sie aber *homogen* sind und die Koeffizienten *nur* die abhängigen Variabeln enthalten (solche Gleichungen nennen wir reduzibel), so kann (2) *durch einfache Vertauschung von* x, y *mit* q, ϑ in ein System von *linearen* Gleichungen mit den unabhängigen Variabeln q, ϑ transformiert werden.

Führt man das *Potential* φ und die *Stromfunktion* ψ ein, wo $u = \frac{\partial \varphi}{\partial x}$, $v = \frac{\partial \varphi}{\partial y}$ und

$$\frac{\partial \varphi}{\partial x} = \frac{1}{\varrho} \frac{\partial \psi}{\partial y}, \qquad \frac{\partial \varphi}{\partial y} = -\frac{1}{\varrho} \frac{\partial \psi}{\partial x}, \tag{9}$$

so stellt (9) wieder ein System nichtlinearer reduzibler Gleichungen dar, da ϱ von $q = \left[\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2\right]^{\frac{1}{2}}$ abhängt. Betrachtet man nun wieder φ und ψ als Funktionen nicht von x, y, sondern von q, ϑ (dies geht auf Chaplygin zurück und entspricht der oben erwähnten Vertauschung), so erhält man die fundamentalen *linearen* Gleichungen

$$\frac{\partial \varphi}{\partial \vartheta} = \frac{q}{\varrho} \frac{\partial \psi}{\partial q}, \quad (M^2 - 1) \frac{\partial \psi}{\partial \vartheta} = \varrho q \frac{\partial \varphi}{\partial q} \tag{10}$$

für φ, ψ als abhängige und q, ϑ als unabhängige Variable.

2. Grenzkurven und Verzweigungskurven. Hilfsformeln. Die Einführung von u, v als neuen unabhängigen und x, y als neuen abhängigen Variablen (Übergang zur Hodographenebene) ist nur möglich, wenn

$$j = \frac{\partial (u, v)}{\partial (x, y)} \neq 0 \tag{11}$$

ist. Man erhält durch diese Hodographentransformation

$$\frac{\partial x}{\partial u} = \frac{1}{j} \frac{\partial v}{\partial y}, \quad \frac{\partial x}{\partial v} = -\frac{1}{j} \frac{\partial u}{\partial y}, \quad \frac{\partial y}{\partial u} = -\frac{1}{j} \frac{\partial v}{\partial x}, \quad \frac{\partial y}{\partial v} = \frac{1}{j} \frac{\partial u}{\partial x}. \tag{12}$$

Damit man anderseits von einer „Hodographenlösung" $\bar{r} = \bar{r}(\bar{q})$ zu einer „physikalischen" Lösung $\bar{q} = \bar{q}(\bar{r})$ zurückkehren kann, muß mit $I j = 1$

$$I = \frac{\partial (x, y)}{\partial (u, v)} \neq 0 \tag{13}$$

sein. Linien, längs denen (13) nicht gilt, werden als *Grenzlinien* der „physikalischen" x, y-Ebene bezeichnet, solche, längs denen (11) versagt, als Verzweigungslinien dieser Ebene[1]).

Wir beginnen mit einigen *Hilfsformeln*. Wir bezeichnen mit ds_1, ds_2 das Linienelement einer ξ- oder η-Linie und setzen

$$h_1 = \frac{ds_1}{d\xi}, \quad h_2 = \frac{ds_2}{d\eta}. \tag{14}$$

[1]) Für eine sog. PRANDTL-MEYER-Strömung ist $j = 0$ in einem zweidimensionalen Gebiet.

Dann folgt für den Vektor $d\bar{r}$ mit den Komponenten dx, dy

$$\frac{\partial \bar{r}}{\partial \xi} = h_1 \bar{\xi}^0, \qquad \frac{\partial \bar{r}}{\partial \eta} = h_2 \bar{\eta}^0, \tag{15}$$

wenn $\bar{\xi}^0$ und $\bar{\eta}^0$ Einheitsvektoren in der $(\vartheta - \alpha)$- und $(\vartheta + \alpha)$-Richtung sind, und mit der Bezeichnung (6):

$$\begin{cases} \dfrac{\partial x}{\partial \xi} = h_1 \cos \varphi^-, & \dfrac{\partial x}{\partial \eta} = h_2 \cos \varphi^+ \\ \dfrac{\partial y}{\partial \xi} = h_1 \sin \varphi^-, & \dfrac{\partial y}{\partial \eta} = h_2 \sin \varphi^+ \end{cases} \tag{15'}$$

und

$$J = \frac{\partial (x, y)}{\partial (\xi, \eta)} = h_1 h_2 \sin 2\alpha. \tag{16}$$

Diese der Abbildung der physikalischen, d.h. der x, y-Ebene auf die charakteristische Ebene entsprechende Funktionaldeterminante werden wir im Überschallgebiet vorwiegend benutzen.

Schreiben wir von nun an kurz $\frac{\partial \varphi}{\partial x} = \varphi_x$, $\frac{\partial \psi}{\partial \xi} = \psi_\xi$ usw., so bestätigt man leicht, daß

$$\begin{cases} \varphi_\xi = h_1 \cdot q \cos \alpha, & \varphi_\eta = h_2 q \cos \alpha \\ \psi_\xi = -h_1 \varrho q \sin \alpha, & \psi_\eta = h_2 \varrho q \sin \alpha \end{cases} \tag{17}$$

ist, und für $D = \frac{\partial (\varphi, \psi)}{\partial (\varrho, \vartheta)}$ findet man (z.B. nach (10) und (7))

$$\begin{cases} D = \dfrac{\operatorname{ctg}^2 \alpha}{\varrho q} (\psi_\vartheta^2 - \psi_Q^2) = -\dfrac{\operatorname{ctg}^2 \alpha}{\varrho q} \psi_\xi \psi_\eta = \dfrac{\varrho}{q} \varphi_\xi \varphi_\eta \\ \qquad = h_1 h_2 \varrho q \cos^2 \alpha = \dfrac{\varrho q}{2} \operatorname{ctg} \alpha \cdot J. \end{cases} \tag{18}$$

Für die in üblicher Weise definierten *Krümmungen* $\varkappa_1, \varkappa_2, \varkappa$ und Krümmungsradien R_1, R_2, R der charakteristischen Kurven und der Stromlinien erhält man

$$\varkappa_1 = \frac{1}{R_1} = \frac{\partial (\vartheta - \alpha)}{\partial s_1} = \frac{\partial (\vartheta - \alpha)}{h_1 \partial \xi}, \quad \varkappa_2 = \frac{1}{R_2} = -\frac{\partial (\vartheta + \alpha)}{\partial s_2} = -\frac{\partial (\vartheta + \alpha)}{h_2 \partial \eta}, \tag{19}$$

$$\varkappa = \frac{\partial \vartheta}{\partial s} = \frac{1}{2 \cos \alpha} \left(\frac{\partial \vartheta}{\partial s_1} + \frac{\partial \vartheta}{\partial s_2} \right) = \frac{1}{2 \cos \alpha} \left(\frac{1}{h_2} - \frac{1}{h_1} \right). \tag{19'}$$

Wir berechnen in gleicher Weise die *Beschleunigung* b und finden:

$$b = \frac{dq}{dt} = q \frac{\partial q}{\partial s} = \frac{q}{2 \cos \alpha} \left(\frac{\partial q}{\partial s_1} + \frac{\partial q}{\partial s_2} \right) = \frac{q^2 \sin \alpha}{2 \cos^2 \alpha} \left(\frac{1}{h_1} + \frac{1}{h_2} \right). \tag{20}$$

Wenn man mit ′ Differentiation nach q bezeichnet und (8) benutzt, so folgt

$$h_1 = -R_1 \left(1 + \frac{\alpha'}{Q'} \right), \quad h_2 = -R_2 \left(1 + \frac{\alpha'}{Q'} \right), \quad h_1/h_2 = R_1/R_2. \tag{21}$$

Setzt man die zweiten gemischten Ableitungen von (15′) einander gleich, und subtrahiert die so erhaltenen Gleichungen, so folgt, nach Vereinfachung, das

System *linearer* Gleichungen für h_1, h_2:

$$(22)\quad \begin{cases} \frac{\partial h_1}{\partial \eta}\sin 2\alpha + \left(\frac{\alpha'}{Q'} - 1\right)(h_2 + h_1 \cos 2\alpha) = 0 \\ \frac{\partial h_2}{\partial \xi}\sin 2\alpha + \left(\frac{\alpha'}{Q'} - 1\right)(h_1 + h_2 \cos 2\alpha) = 0, \end{cases}$$

die für die folgende Diskussion benutzt werden.

3. Fall der Unterschallgeschwindigkeit. Wenn $M<1$ ist in

$$(18')\qquad D = \frac{\partial(\varphi, \psi)}{\partial(q, v)} = -\frac{1}{\varrho q}\left[q^2 \psi_q^2 + (1 - M^2)\psi_\vartheta^2\right],$$

so verschwindet D nur, wenn beide Summanden in der Klammer verschwinden, d.h. wenn

$$q\psi_q = 0 \quad \text{und} \quad \psi_\vartheta = 0,$$

und nach (10)

$$\varphi_\vartheta = 0, \qquad q\varphi_q = 0.$$

Also ist dann entweder $q=0$ (Stagnationspunkt) und außerdem $\psi_\vartheta = \varphi_\vartheta = 0$; oder es ist $q \neq 0$, dann *müssen alle vier Ableitungen* $\psi_q, \psi_\vartheta, \varphi_q, \varphi_\vartheta$ *und daher auch* $x_q, x_\vartheta, y_q, y_\vartheta$ *verschwinden.* (Man kann im letzteren Falle leicht sehen, daß der singuläre Punkt ein *Sattelpunkt von* $\varphi(q, \vartheta)$ *und* $\psi(q, \vartheta)$ *ist, wenn nicht alle Ableitungen zweiter Ordnung verschwinden.*)

Ein ganz ähnliches Ergebnis gilt, falls $M<1$, für Verzweigungssingularitäten. Doch gehen wir auf diese nicht ein.

4. Grenzlinien L_1 und L_2. In der Diskussion des Verschwindens von J ist es offenbar von Vorteil, den Ausdruck für J gemäß (16) in seine Bestandteile zu zerlegen. Dementsprechend betrachten wir zu einer gegebenen Strömung den Ort $h_1(\xi, \eta) = 0$ in der ξ, η-Ebene und seine Abbildung auf die x, y-Ebene. Wir nehmen an, daß α weder 0° noch 90° ist. Sei nun p_0 ein Punkt mit Koordinaten ξ_0, η_0, für den $h_1(\xi_0, \eta_0) = 0$, $\left(\frac{\partial h_1}{\partial \eta}\right)_{\xi_0, \eta_0} \neq 0$ ist. Dann gibt es eine Kurve $\eta = g(\xi)$, $\eta_0 = g(\xi_0)$ längs der $h_1(\xi, \eta) = 0$. Wir nennen sie l_1, und ihr Bild L_1 in der x, y-Ebene *Grenzlinie*. (Natürlich ist l_1 selbst Verzweigungslinie der ξ, η-Ebene. Doch beziehen wir uns in der Namengebung nur auf die x, y-Ebene.) — Ein Punkt P der x, y-Ebene, für dessen Bild p die Gleichung $h_1(\xi, \eta) = 0$ erfüllt ist, heißt *Grenzpunkt.*

Sei m ein Punkt von l_1 und dort sowohl $\frac{\partial h_1}{\partial \xi}$ wie $\frac{\partial h_1}{\partial \eta}$ ungleich Null. Daher ist nach (22) $h_2 \neq 0$. Also hat nach (15) L_1 in M, dem Bilde von m die η-Richtung. Dasselbe gilt für jede Kurve C durch M, deren Bild c durch m etwa $\eta = f(\xi)$, dort *nicht* die ξ-Richtung hat, die wir als *Ausnahmerichtung* bezeichnen. Alle solchen Kurven, insbesondere die C^+ durch M, berühren die L_1 in M. Das gilt auch im Besonderen von den *Kurven konstanter Geschwindigkeit* q *und den Kurven konstanter Richtung* ϑ. Denn z.B. für letztere ist $d\vartheta = 0$, aber wegen $\alpha \neq 90°$ und (16): $d\eta = \frac{1}{2}(dQ + d\vartheta) = \frac{1}{2} dQ \neq 0$. Daher

berühren alle diese Kurven die Kurve L_1 in M und haben dort in jedem Punkt die C^+-Richtung.

Hat aber c in m die Ausnahmerichtung, so ist $\frac{d\eta}{d\xi} = f'(\xi) = 0$. Man kann nicht mehr schließen, daß die entsprechende C in M die η-Richtung hat, da nun $d\xi$ und $d\eta$ von gleicher Ordnung klein sind. Man sieht dann leicht, daß in M sowohl $\frac{dx}{d\xi} = \frac{\partial x}{\partial \xi} + \frac{\partial x}{\partial \eta} f'(\xi)$ wie auch $\frac{dy}{d\xi}$ verschwinden, während $\frac{d^2x}{d\xi^2}$ und $\frac{d^2y}{d\xi^2}$ nicht beide verschwinden, und daß folglich *C in M eine Spitze hat*. Insbesondere hat daher die Kurve C^- in M eine Spitze. Ebenso haben *Stromlinien* und *Potentiallinien* jede eine Spitze in M. Denn nach (17) folgt aus $h_1 = 0$, daß $\psi_\xi = 0$ ist, und da $d\psi = \psi_\xi\, d\xi + \psi_\eta\, d\eta = \psi_\eta\, d\eta = 0$, und da $\psi_\eta \neq 0$, wegen $h_2 \neq 0$, ist $d\eta = 0$ in m. Daher hat das Bild der

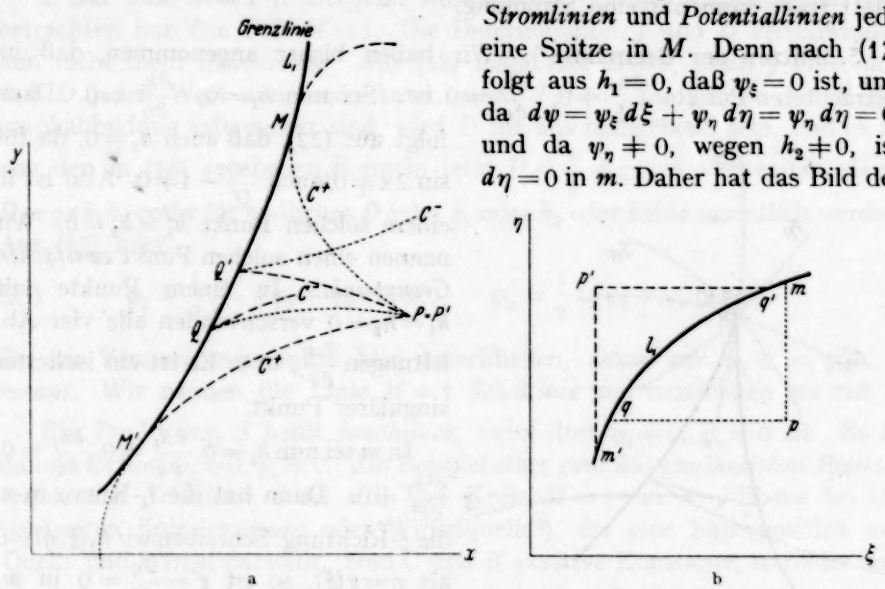


Fig. 1a u. b. Umgebung der Grenzlinie in x, y-Ebene und ξ, η-Ebene

Stromlinie in m die Ausnahmerichtung und die Stromlinie in M eine Spitze. Dies ist in Übereinstimmung mit (20), wonach $\varkappa$ an der Grenzlinie unendlich wird. Nach (20) *wird an einer Grenzlinie auch die Beschleunigung unendlich.*

Wir fassen zusammen: *Wir betrachten in der ξ, η-Ebene den Ort der Punkte $h_1(\xi, \eta) = 0$, im allgemeinen eine Kurve l_1; wir betrachten Punkte m auf l_1, für die $\frac{\partial h_1}{\partial \xi} \neq 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$ ist. Das Bild L_1 von l_1, die Grenzlinie, ist in allen solchen Punkten Einhüllende von C^+-Charakteristiken, von Kurven konstanter Richtung und Kurven konstanter Geschwindigkeit. Die C^--Charakteristiken, die Stromlinien und die Potentiallinien haben Spitzen auf L_1.* Für $h_2 = 0$ und die entsprechende Grenzlinie L_2 sind die Rollen von ξ und η vertauscht.

An der Grenzlinie stoßen zwei Lösungen zusammen (Fig. 1). Betrachten wir die Korrespondenz der beiden Ebenen x, y und ξ, η in der Umgebung dieser Linie. Ein Punkt P in der x, y-Ebene ist das Bild zweier Punkte p und p'. Durch P gehen zwei C^+-Charakteristiken, die die L_1 berühren, und zwar in M und M', mit entsprechenden Punkten m und m' auf l_1. Auf der

η-Linie durch m liegt p, auf der durch m' liegt p'. Diese beiden Punkte liegen auf entgegengesetzten Seiten der Kurve l_1, da η entlang PM und PM' in entgegengesetztem Sinne variiert. Durch P gehen auch zwei C^--Charakteristiken PQ und PQ', die mit den Charakteristiken PM und PM' die Winkel $2\alpha_1$ und $2\alpha_2$ einschließen. Der Punkt q liegt auf der ξ-Linie durch p und auf l_1, und q' liegt auf der ξ-Linie durch p' und auf l_1.

Wir erhalten in üblicher Weise eine eineindeutige Abbildung, wenn wir jeder der beiden Seiten der l_1 ein Blatt der physikalischen Ebene zuordnen (in der Fig. 1 mittels verschiedener Stricharten angedeutet). Jedes solche Blatt trägt dann nur eine Strömung.

5. Spitzen der Grenzlinie[2]). Wir haben bisher angenommen, daß im betrachteten Punkte $\frac{\partial h_1}{\partial \xi} \neq 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$ ist. Sei nun $h_1 = 0$, $\frac{\partial h_1}{\partial \eta} = 0$. Dann folgt aus (22), daß auch $h_2 = 0$, da für $\sin 2\alpha \neq 0$ auch $\frac{\alpha'}{Q'} - 1 \neq 0$. Also ist in einem solchen Punkt $h_1 = h_2 = 0$. Wir nennen einen solchen Punkt *zweifachen Grenzpunkt*. In einem Punkte mit $h_1 = h_2 = 0$ verschwinden alle vier Ableitungen $\frac{\partial x}{\partial \xi}$, usw. Er ist ein isolierter singulärer Punkt.

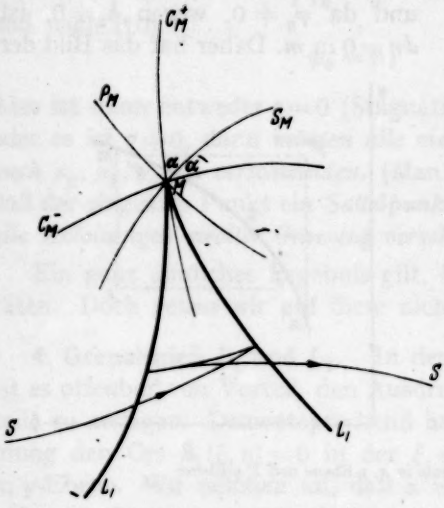


Fig. 2. Grenzlinie mit Spitze in M mit Charakteristiken C_M^+, C_M^-, Stromlinie S_M und Potentiallinie P_M in M. Stromlinie S mit Spitzen auf L_1

In m sei nun $h_1 = 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$, $\frac{\partial h_1}{\partial \xi} = 0$, $\frac{\partial^2 h_1}{\partial \xi^2} \neq 0$. Dann hat die l_1-Kurve in m die ξ-Richtung. Schreiben wir $h_1(\xi, \eta) = 0$ als $\eta = g(\xi)$, so ist $g' = \frac{d\eta}{d\xi} = 0$ in m, aber $g'' \neq 0$ in m. Man zeigt dann, daß in m zwar $\frac{dx}{d\xi} = \frac{dy}{d\xi} = 0$ ist, daß aber $\frac{d^2x}{d\xi^2}$ und $\frac{d^2y}{d\xi^2}$ nicht beide verschwinden, *und daher L_1 in M eine Spitze hat.* Ferner sieht man, daß *die Kurve C^+ in M die Spitzentangente berührt, ohne selbst dort eine Spitze zu haben* (Fig. 2). Dasselbe gilt für jede Kurve durch M, deren Bild in m nicht die Ausnahmerichtung hat, also z. B. für die Kurven konstanter Richtung — damit meinen wir Kurven längs deren ϑ einen konstanten Wert hat — und die Kurven konstanter Geschwindigkeit. Die Kurve C^-, deren Bild die Ausnahmerichtung hat, schließt mit der Spitzentangente in M den Winkel 2α ein. Stromlinie und Potentiallinie in M halbieren diesen Winkel von innen und von außen. Die Bildkurve l_1 hat in m *ein Extremum*. Sie wird von der Stromlinie durch m in der ξ-Richtung *berührt*.

[2]) Dieser Teil fehlt ganz in den unter [4] und [7] zitierten Werken. In [3] und [5] sind die Resultate mit anderer Methode hergeleitet.

Diesen Spitzen der Grenzlinie, Extrema von l_1, kommt eine wichtige Rolle zu: Die Stromlinie, die durch die Spitze von L_1 geht, trennt diejenigen Stromlinien, die die Kurve L_1 nicht schneiden und völlig glatt verlaufen, von denen, die L_1 schneiden und daher an den Ästen von L_1 Spitzen haben (Fig. 2). Das Gebiet zwischen den Ästen von L_1 ist dreifach bedeckt. All dies kann z.B. an der bekannten RINGLEB-Strömung [1] verfolgt werden. Auch bei Fragen, die in Verbindung mit der heute so viel diskutierten „gemischten" Strömung um ein Profil auftreten, spielen die Spitzen der Grenzlinien eine Rolle.

6. Der Fall $M=1$[3]. Bis jetzt wurde $\alpha \neq 90°$, $M \neq 1$ angenommen. Wir betrachten nun den Fall $M=1$. Die Determinanten J und D verschwinden nun nicht mehr gleichzeitig. Aus (18) folgt, daß $D=0$, wenn $J<\infty$, und daß D jeden Wert annehmen kann, wenn $J\to\infty$. Da wir an der Hodographabbildung interessiert sind, wird D für uns maßgebend sein, und es ist von den in (18) gegebenen Formeln jetzt $D=\frac{\varrho}{q}\varphi_\xi\varphi_\eta$ zu verwenden, da in $D=\varrho q h_1 h_2 \cos^2\alpha$ für endliches D jetzt h_1 oder h_2 oder beide unendlich werden. Aus (17) folgt

$$(23)\qquad \varphi_\xi=\frac{1}{\varrho}\left[-q\psi_q+\psi_\vartheta \operatorname{ctg}\alpha\right],\qquad \varphi_\eta=\frac{1}{\varrho}\left[q\psi_q+\psi_\vartheta\operatorname{ctg}\alpha\right].$$

Da wir Verzweigungspunkte hier ausschließen, *setzen wir* $\varphi_\xi<\infty$, $\varphi_\eta<\infty$ *voraus*. Wir nennen die Linie $M=1$ *Schallinie* und bezeichnen sie mit S.

Ein Punkt von S heißt *gewöhnlich*, wenn dort $\varphi_\xi\neq 0$, $\varphi_\eta\neq 0$ ist. Es ist dann a fortiori $\psi_\xi\neq 0$, $\psi_\eta\neq 0$. Ein Beispiel einer *ganz aus gewöhnlichen Punkten bestehenden Schallinie* S ist z. B. der Kreis $M=1$ der x, y-Ebene bei der bekannten *Spiralströmung* oder Wirbelquelle[4], die eine Superposition von Quelle und Wirbel darstellt. Sind C und K positive Konstante, so findet man

$$(24)\qquad \begin{cases} \varphi_\xi=K\cdot\frac{1}{\varrho}\operatorname{ctg}\alpha-C, & \varphi_\eta=K\cdot\frac{1}{\varrho}\operatorname{ctg}\alpha+C \\ \psi_\xi=-K+C\varrho\operatorname{tang}\alpha, & \psi_\eta=K+C\varrho\operatorname{tang}\alpha \\ \psi_q=C\frac{\varrho}{q}, \quad \psi_\vartheta=K, & b=\frac{K\varrho q^3}{K^2\operatorname{ctg}^2\alpha-C^2\varrho^2}. \end{cases}$$

Es ist also für $M=1$: $\psi_\vartheta\neq 0$, $\psi_q\neq 0$, $\varphi_\xi=-C$, $\varphi_\eta=C$, also alle Punkte gewöhnlich. Die Stromlinien beider Strömungen schneiden den Kreis $M=1$ unter einem von einem Rechten verschiedenen Winkel; *die Charakteristiken berühren diesen Kreis S daher nicht.*

An dem *Schallpunkt einer* L_1: $h_1=0$, ist $\psi_\xi=0$, und daher a fortiori $\varphi_\xi=0$, und analoges gilt für die L_2. Es kann gezeigt werden, daß ein Kurvenbogen L_1 oder L_2, längs dem überall $M=1$ ist, nicht möglich ist.

[3] Diese Überlegungen sind im wesentlichen neu. Herrn G. S. S. LUDFORD bin ich für wertvollen Rat zu Dank verpflichtet.

[4] Die hier benutzten Beispiele mögen in jedem der üblichen Lehrbücher, z.B. in dem bekannten Buche von SAUER, oder in den unter [4] oder [5] zitierten nachgelesen werden.

Eine neue Art von Grenzpunkt, wir nennen ihn *Schallgrenzpunkt*, typisch für $\alpha = 90°$, ist durch $\varphi_\xi = 0$, $\varphi_\eta = 0$, $\psi_\xi \neq 0$, $\psi_\eta \neq 0$, daher $h_1 \neq 0$, $h_2 \neq 0$ charakterisiert. Aus den ersten zwei Bedingungen folgt $\psi_q = 0$, $\psi_\vartheta \operatorname{ctg} \alpha = 0$. Also entweder $\psi_q = 0$, $\psi_\vartheta = 0$, oder $\psi_q = 0$, $\psi_\vartheta \neq 0$, aber so, daß $\psi_\vartheta \operatorname{ctg} \alpha \to 0$ mit $\alpha \to 90°$. Im ersten Fall handelt es sich um einen isolierten Ausnahmepunkt [vgl. (10) und (15′)]. Wichtiger ist der zweite Fall, da dieser offenbar entlang einer ganzen Kurve zutreffen kann, die wir eine *Schallgrenzkurve* L_t nennen. Das einfachste Beispiel einer Kurve L_t ist die *Radialströmung* (Quelle). Aus (24) für $C = 0$ ergibt sich, daß *für den ganzen Kreis* $M = 1$

$$\varphi_\eta = 0, \quad \varphi_\xi = 0, \quad D = 0, \quad b \to \infty, \quad h_1 \neq 0, \quad h_2 \neq 0. \tag{25}$$

In diesem Beispiel gibt es keine andere Grenzlinie, da h_1 und h_2 nach (17) überall endlich sind.

Gemäß unserer Definition ist jede Schallgrenzlinie eine Äquipotentiallinie. Daher ist die Stromlinienrichtung überall normal auf L_t, und da $\alpha = 90°$, so folgt, daß *sowohl die* C^+ *wie die* C^- *die Kurve* L_t *berühren*.

Kehren wir zu unserem Beispiel zurück, so sehen wir, daß *die Linien konstanter Richtung* ϑ *nun die* L_t *nicht berühren* (im Gegensatz zu unseren früheren Resultaten für die L_1 und L_2). Diese Linien sind nämlich die Radien, die auf der L_t senkrecht stehen und *mit den Stromlinien zusammenfallen*, und zwar mit den Stromlinien *beider Strömungen*. Jeder Punkt des Schallkreises gilt daher als ausgeartete Spitze.

Die Linie konstanter Richtung wird mit der Stromlinie lokal immer zusammenfallen, wo $\psi_q = 0$, $\psi_\vartheta \neq 0$. Denn aus $d\psi = \psi_q\, dq + \psi_\vartheta\, d\vartheta = \psi_\vartheta\, d\vartheta = 0$ folgt $d\vartheta = 0$. Dies kann, wie wir gesehen haben, an einer L_t zutreffen: $\psi_q = 0$, $\psi_\vartheta \neq 0$. Es kann auch für den Schallpunkt einer L_1 zutreffen, denn für $\alpha = 90°$ kann $\psi_\xi = -\psi_\vartheta + q\,\psi_q \operatorname{tang} \alpha$ verschwinden, wenn $\psi_q = 0$, $\psi_\vartheta \neq 0$. Dieselbe Formel zeigt aber, daß dies für $\alpha \neq 90°$ nicht möglich ist. Und in der Tat haben wir in **4** und **5** gefunden, daß an einer Kurve L_1 die Stromlinie mit der Linie konstanter Richtung den Winkel α macht.

Die aus der Definition von φ und ψ, gemäß (9), folgende Beziehung

$$dx = \frac{1}{q}\left(\cos\vartheta\, d\varphi - \frac{1}{\varrho}\sin\vartheta\, d\psi\right), \qquad dy = \frac{1}{q}\left(\sin\vartheta\, d\varphi + \frac{1}{\varrho}\cos\vartheta\, d\psi\right)$$

ergibt an der L_t, wo $d\varphi = 0$, $\psi_q = 0$

$$dx = -\frac{1}{\varrho q}\psi_\vartheta \sin\vartheta\, d\vartheta, \qquad dy = \frac{1}{\varrho q}\psi_\vartheta \cos\vartheta\, d\vartheta$$

$$\frac{dy}{dx} = -\cot\vartheta, \qquad \text{wenn} \quad d\vartheta \neq 0.$$

Das heißt, daß jedes Hodographelement mit $d\vartheta \neq 0$ in ein Element der x, y-Ebene senkrecht zur Stromrichtung abgebildet wird. Dies gilt, wie schon früher gesehen, für die L_t selbst, die eine Kurve konstanter Geschwindigkeit q und eine Potentiallinie ist, und für die beiden Charakteristiken C^- und C^+, während Stromlinie und Isokline, $d\vartheta = 0$, senkrecht auf der L_t stehen, da sie

im Hodograph die radiale Ausnahmerichtung aufweisen. Eine kleine Rechnung zeigt, daß solche Ausnahmekurven an der L_i *Spitzen* haben. Die Charakteristiken C^- und C^+ liegen auf verschiedenen Seiten der Isokline (bilden also miteinander einen Winkel von 180°), da ihre Bilder im Hodograph durch die Radien getrennt werden. Wir fassen zusammen: *Die gemäß 1.d) definierte Schallgrenzkurve, die eine Equipotentiallinie und eine Linie konstanter Geschwindigkeit ist, ist Einhüllende beider Scharen von Charakteristiken, und Ort der Spitzen von Stromlinien und Linien konstanter Richtung, die auf ihr senkrecht stehen.*

Da an einem Punkt mit $\psi_\xi = 0$ auch φ_ξ verschwindet, erkennt man, daß die folgende Klassifikation vollständig ist:

1.	$\psi_\xi \neq 0,\ \psi_\eta \neq 0$	a) $\varphi_\xi \neq 0,\ \varphi_\eta \neq 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta \neq 0$
		c) $\varphi_\xi \neq 0,\ \varphi_\eta = 0$
		d) $\varphi_\xi = 0,\ \varphi_\eta = 0$
2.	$\psi_\xi = 0,\ \psi_\eta \neq 0$	a) $\varphi_\xi = 0,\ \varphi_\eta \neq 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta = 0$
3.	$\psi_\xi \neq 0,\ \psi_\eta = 0$	a) $\varphi_\xi \neq 0,\ \varphi_\eta = 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta = 0$
4.	$\psi_\xi = 0,\ \psi_\eta = 0$	$\varphi_\xi = 0,\ \varphi_\eta = 0$.

Hier charakterisiert 1.a) den *gewöhnlichen Punkt*, 1.d) *den Schallgrenzpunkt*, 2. und 3. den *Schallpunkt einer* L_1 *oder einer* L_2, 4. den *zweifachen Grenzpunkt*.

Wie früher bemerkt, können die Fälle 2., 3., 4. nur in isolierten Punkten auftreten. Dasselbe kann von 1.b) und 1.c) leicht eingesehen werden, so daß in der Tat die Schallgrenzlinie als einzige Grenz*linie* für $M = 1$ erscheint.

Wir betrachten noch die *Beschleunigung*

$$b = \frac{1}{2} q^3 \operatorname{tang} \alpha \cdot \left(\frac{1}{\varphi_\xi} + \frac{1}{\varphi_\eta} \right). \tag{20'}$$

Es geht also b an einer L_1, *an einer* L_2, *und an einer* L_i *gegen unendlich.* An einem regulären Punkt von S kann man keine Aussage machen da $\varphi_\xi + \varphi_\eta = 0$ sein kann.

Ein interessantes Beispiel für den *Schallpunkt einer* L_1 (und einer L_2), bietet die erwähnte RINGLEB-Strömung. Wir haben in diesem Falle (k eine Konstante):

$$\begin{cases} \psi_\xi = -\dfrac{k}{q} (\sin \vartheta \operatorname{tang} \alpha + \cos \vartheta), & \psi_\eta = -\dfrac{k}{q} (\sin \vartheta \operatorname{tang} \alpha - \cos \vartheta) \\ \varphi_\xi = \dfrac{k}{\varrho q} (\sin \vartheta + \cos \vartheta \operatorname{ctg} \alpha), & \varphi_\eta = -\dfrac{k}{\varrho q} (\sin \vartheta - \cos \vartheta \operatorname{ctg} \alpha), \end{cases} \tag{26}$$

und die Hodographengleichung der Grenzlinie lautet

$$\cos \vartheta = \pm \frac{1}{M} = \pm \sin \alpha.$$

Die Grenzlinie besteht aus zwei Teilen, einem oberhalb der x-Achse, entsprechend dem Ast $\cos\vartheta - 1/M = 0$, oder $\vartheta = 90° - \alpha$ $(0 \leq \vartheta \leq 90°)$; der andere unterhalb der x-Achse, entsprechend $\cos\vartheta + 1/M = 0$, oder $\vartheta = 90° + \alpha$ $(90° \leq \vartheta \leq 180°)$. Auf dem oberen ist für $\alpha = 90°$, $\psi_\xi \neq 0$, $\psi_\eta = 0$, daher ist dies eine L_2, deren Schallpunkt die Hodographkoordinaten $q = a$, $\vartheta = 0$ besitzt. Entlang dem anderen Ast ist für $\alpha = 90°$, $\psi_\xi = 0$, $\psi_\eta \neq 0$, also ist dies eine L_1, deren Schallpunkt die Koordinaten $q = a$, $\vartheta = 180°$ besitzt. [Für jeden dieser *beiden* Schallpunkte gilt $\varphi_\xi = 0$, $\varphi_\eta = 0$, also liegen die Fälle 2.b) und 3.b) obiger Klassifikation vor.] Diesen beiden verschiedenen Hodographpunkten entspricht *derselbe* Punkt in der x, y-Ebene, nämlich der Schnittpunkt A der Grenzlinie mit der x-Achse. Dieser Punkt ist *nicht ein zweifacher Grenzpunkt*, da er zwei verschiedenen Hodographenpunkten entspricht. Für jeden dieser beiden Punkte folgt, wegen $\varphi_\xi = 0$, $\varphi_\eta = 0$, $\psi_q = 0$, $\psi_\vartheta \neq 0$, aus dem früheren, daß die Grenzlinie in A beide Charakteristiken berührt, und daß die Richtung konstanten ϑ mit der Stromlinienrichtung zusammenfällt, was auch beides der Fall ist.

Die beiden Äste L_1 und L_2 schneiden sich im Unendlichen. *Dort* ist für $\alpha = 0°$, $\vartheta = 90°$ sowohl $h_1 = 0$ wie $h_2 = 0$, also ein zweifacher Grenzpunkt.

Auf weitergehende Diskussion von Grenzpunkten oder Grenzlinien gehen wir hier nicht ein. Wir wollten in dieser kleinen Arbeit die *Methode der Untersuchung* darlegen und zur Herleitung einiger Hauptresultate verwenden. Aus Raummangel muß auch eine Diskussion der Verzweigungslinien unterbleiben.

Literatur

[1] RINGLEB, F.: Exakte Lösungen der Differentialgleichungen einer adiabatischen Gasströmung. Z. angew. Math. Mech. **20**, 185—198 (1940). — [2] TOLLMIEN, W.: Grenzlinien adiabatischer Potentialströmungen. Z. angew. Math. Mech. **21**, 140—152 (1941). — [3] CRAGGS, J. W.: The breakdown of the hodograph transformation for irrotational compressible fluid flow in two dimensions. Proc. Cambridge Phil. Soc. **44**, 360—379 (1948). — [4] COURANT, R., and K. O. FRIEDRICHS: Supersonic Flow and Shock Waves. New York: Interscience Publishers, Inc. 1948. — [5] Modern Development in Fluid Dynamics High Speed Flow (particularly Chapter VII.) Edited by L. Howarth. Oxford University-Press 1953. — [6] MEYER, R. E.: Focussing effects in twodimensional supersonic flow. Phil. Trans. Roy. Soc. Lond. **242**, 153—171 (1949). — [7] MANDEL, J.: Equilibres par tranches planes des solides à la limite d'écoulement. Paris: Louis Jean 1942. — [8] MISES, R. v.: Three remarks on the theory of the ideal plastic body. H. Reissner Anniversary Volume, p. 415—429. Edwards Brothers. Ann Arbor, Mich. 1949. — [9] GEIRINGER, H.: Some recent results in the theory of an ideal plastic body. Advances in Applied Mechanics, Vol. III, p. 197—294. New York: Academic Press 1953.

Harvard University, Cambridge (U.S.A.)

(Eingegangen am 3. Oktober 1955)

MORDELL, L. J.
Math. Zeitschr. Bd. 63, S. 525—528 (1956)

The Minimum of an Inhomogeneous Quadratic Polynomial in n Variables

Dedicated to the memory of I. SCHUR

By
L. J. MORDELL

Let

$$f(x) = f(x_1, x_2, \ldots, x_n) \tag{1}$$

be a real function of n real variables $(x) = (x_1, x_2, \ldots, x_n)$, and let (x') be a given set of real numbers. A well known problem is to find a number k independent of (x') such that there exist real numbers (x) with $(x) \equiv (x')$ (mod 1) and

$$f(x) \leq k. \tag{2}$$

An alternative form of the question is to consider integers (x) such that

$$f(x + x') \leq k. \tag{3}$$

Put in this form, we see that $f(x)$ may be replaced by an equivalent form $f(x) = g(y)$ arising from the unimodular integral substitution

$$x_l = \sum_{m=1}^{n} \lambda_{lm} y_m \qquad (l = 1, 2, \ldots, n).$$

Thus if we define (y') by

$$x_l' = \sum_{m=1}^{n} \lambda_{lm} y_m',$$

the problem of making $f(x + x') \leq k$ is equivalent to making $g(y + y') \leq k$.

Let now

$$f(x) = \sum_{r,s=1}^{n} a_{rs} x_r x_s + 2 \sum_{r=1}^{n} a_r x_r, \tag{4}$$

where the coefficients are real, and $a_{rs} = a_{sr}$, $a_{rr} > 0$ for $1 \leq r, s \leq n$, and the determinant

$$A = \|a_{rs}\| \neq 0.$$

In the particular case when $n = 2$, and $a_1 = a_2 = 0$ and $f(x)$ is positive definite, DIRICHLET, in a paper[1]) written in 1848, has shown that

$$4(a_{11} a_{22} - a_{12}^2)\, k = a_{11} a_{22} (a_{11} - 2 a_{12} + a_{22}) \tag{5}$$

[1]) Werke Bd. II, S. 29—48.

when $f(x)$ is replaced by the reduced equivalent form for which

$$a_{12} \geq 0, \quad a_{11} \geq a_{12}, \quad a_{22} \geq a_{12}, \tag{6}$$

and that this is the best possible result.

I am not aware that any similar results have been found for $n \geq 3$. By means of an exceedingly simple idea, I can include Dirichlet's result in the

Theorem. *Let k be defined by the determinantal equation*

$$\begin{vmatrix} a_{11}, & a_{12}, & \ldots, & a_{1n}, & \frac{1}{2}a_{11} + a_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}, & a_{n2}, & \ldots, & a_{nn}, & \frac{1}{2}a_{nn} + a_n \\ -\frac{1}{2}a_{11} + a_1, & & \ldots, & -\frac{1}{2}a_{nn} + a_n, & k \end{vmatrix} = 0 \tag{7}$$

Then if the variables (x) can be divided into two sets of variables $(x_\lambda, x_{\lambda'}, \ldots)$ and $(x_\mu, x_{\mu'}, \ldots)$ such that all the cross products $a_{\lambda\lambda'}$ $(\lambda \neq \lambda')$, and $a_{\mu\mu'}$ $(\mu \neq \mu')$, are ≤ 0, then numbers (x) with assigned residues (mod 1) *exist such that*

$$f(x) \leq k. \tag{8}$$

There is of course no real loss of generality in taking $a_r = 0$, and then the result takes the simpler form now given.

Let

$$F(x) = \sum_{r,s=1}^{n} A_{rs} x_r x_s,$$

where A_{rs} denotes the corresponding cofactor of a_{rs} in A. Then

$$4Ak = F(a_{11}, a_{22}, \ldots, a_{nn}), \tag{9}$$

since now (7) becomes

$$\begin{vmatrix} a_{11}, & a_{12}, & \ldots, & a_{1n}, & a_{11} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}, & & \ldots, & a_{nn}, & a_{nn} \\ a_{11}, & & \ldots, & a_{nn}, & -4k \end{vmatrix} = 0.$$

When $n = 2$, (9) is the best possible result if and only if the conditions (6) hold, and then $A \geq 0$.

We show first that on replacing (x) by $(x + \xi)$ for appropppriate real (ξ), the inequality (8) takes the form

$$g(x) = \sum_{r=1}^{n} a_{rr}(x_r^2 - x_r) + \sum_{r \neq s} a_{rs} x_r x_s \leq 0. \tag{10}$$

The condition for this is, on noting the new coefficients of x_r,

$$\begin{aligned} &a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n + a_1 + \tfrac{1}{2}a_{11} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{n1}\xi_1 + \cdots + a_{nn}\xi_n + a_n + \tfrac{1}{2}a_{nn} = 0, \\ &f(\xi_1, \xi_2, \ldots, \xi_n) = k. \end{aligned}$$

Multiply the first n equations in order by $-\xi_1, -\xi_2, \ldots, -\xi_n$, and add to the last. Then

$$(a_1 - \tfrac{1}{2}a_{11})\xi_1 + \cdots + (a_n - \tfrac{1}{2}a_{nn})\xi_n - k = 0.$$

Eliminating $\xi_1, \xi_2, \ldots, \xi_n$ from these equations, we have the equation (7) for k.

We show now that with the assigned conditions on the signs of the a, (10) is solvable with any assigned residues (mod 1) for (x). Clearly, we may without loss of generality suppose that $0 \leq x_r < 1$ for $r = 1, 2, \ldots, n$; and we note equality occurs in (10) when $x_r = 0$. We prove that (10) is satisfied by at least one of the sets $(x_r - \varepsilon)$ where $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ and $\varepsilon_r = 0$ or 1. If this is not so, then for these 2^n sets of values, $g(x) > 0$. Take in particular

$$g(x_1, x_2, \ldots, x_n) > 0, \quad g(x_1 - 1, x_2, \ldots, x_n) > 0. \tag{11}$$

We now apply an averaging process for x_1. Multiply the two inequalities in (11) by $1 - x_1$, x_1 respectively and add. Note that

$$x^2(1-x) + x(x-1)^2 = x(1-x), \quad x(1-x) + (x-1)x = 0.$$

Write $X = x - x^2$. Then

$$a_{11}X_1 - a_{22}X_2 - \cdots - a_{nn}X_n + \sum_{r,s \neq 1}^{r \neq s} a_{rs}x_r x_s > 0. \tag{12}$$

So averaging for x_2,

$$-a_{11}X_1 + a_{22}X_2 - \cdots - a_{nn}X_n + \sum_{r,s \neq 2}^{r \neq s} a_{rs}x_r x_s > 0. \tag{13}$$

Suppose $n = 2$, then by addition, we have the contradiction $0 > 0$.

We prove now that (10) gives the best possible result, i.e. that $g(x) \geq 0$ for all integers (x) if and only if

$$a_{12} \geq 0, \quad a_{11} \geq a_{12}, \quad a_{22} \geq a_{12}. \tag{14}$$

These conditions are obviously necessary from $g(1, 1) \geq 0$, $g(1, -1) \geq 0$, $g(-1, 1) \geq 0$. Next, they are sufficient. For

$$\begin{aligned} a_{11}(x_1^2 - x_1) + 2a_{12}x_1x_2 + a_{22}(x_2^2 - x_2) &\geq a_{12}(x_1^2 - x_1 + 2x_1x_2 + x_2^2 - x_2) \\ &\geq a_{12}(x_1 + x_2)(x_1 + x_2 - 1) \\ &\geq 0, \end{aligned}$$

and the result follows.

The conditions (14) mean that here $A \geqq 0$ and that $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ is a reduced form when $a_{12} \leqq \frac{1}{2}a_{11}$, and that

$$a_{11}(x_1 - x_2)^2 + 2a_{12}(x_1 - x_2)x_2 + a_{22}x_2^2$$

is when $a_{12} > \frac{1}{2}a_{11}$. This gives Dirichlet's theorem.

Suppose now $n \geqq 3$. We may suppose that $a_{rs} \leqq 0$, $r \neq s$, for $r \leqq m$, $s \leqq m$, say, and also $a_{rs} \leqq 0$ for $r > m$, $s > m$, $r \neq s$.

On averaging (12) for the variable $x_2, x_3, \ldots, x_m$ successively, we obtain

$$\sum_{r \leqq m} a_{rr}X_r - \sum_{r > m} a_{rr}X_r + \sum_{r,s \neq 1,2,\ldots m}^{r \neq s} a_{rs}x_r x_s > 0. \tag{15}$$

So on averaging for the variables $x_{m+1}, \ldots, x_n$, we have

$$\sum_{r > m} a_{rr}X_r - \sum_{r \leqq m} a_{rr}X_r + \sum_{r,s \neq m+1,\ldots n}^{r \neq s} a_{rs}x_r x_s > 0. \tag{16}$$

On adding (15), (16), we have a contradiction since the X's cancel and the coefficients occurring satisfy $a_{rs} \leqq 0$.

We note that when $n = 3$, the result holds if $a_{12} \leqq 0$ (or if $a_{23} \leqq 0$ or if $a_{31} \leqq 0$), and that when $n = 4$, the result holds if $a_{12} \leqq 0$, $a_{34} \leqq 0$.

The result, however, is not best possible since (14) shows that a necessary condition is obviously $a_{rs} \geqq 0$, $r, s = 1, \ldots, n$.

St. John's College, Cambridge, England

(Eingegangen am 13. Juni 1955)

BRAUER, A.
Math. Zeitschr. Bd. 63, S. 529—541 (1956)

On the Schnirelmann density of the sum of two sequences

In memory of my teacher and friend I. Schur

By

ALFRED BRAUER

In 1936, SCHUR published his paper "Über den Begriff der Dichte in der additiven Zahlentheorie" [18]. Under the influence of this paper and discussions with SCHUR I wrote my papers [2] and [3]. The result of [3] will be improved by the present paper.

Let A be a sequence of positive integers $a_1 < a_2 < \cdots$ and $A(n)$ the number of elements $a_\nu \leq n$. The greatest lower bound of the quotients $A(n)/n$ is called the SCHNIRELMANN density α of A [17].

The sum $A + B$ of the sequences A and B is defined as the sequence of all integers of the form

$$e_1 a + e_2 b \quad (a \in A,\ b \in B;\ e_1 = 0 \text{ or } 1,\ e_2 = 0 \text{ or } 1).$$

We define the "modified" sum $A \dotplus B$ as the sequence of all integers of the form

$$a + e_2 b \quad (a \in A,\ b \in B;\ e_2 = 0 \text{ or } 1). \tag{1}$$

In the following, we denote the density of $A + B$ by γ, that of $A \dotplus B$ by γ', and that of B by β. It was conjectured by LANDAU [9, 10] that

$$\gamma \geq \min(1, \alpha + \beta). \tag{2}$$

This conjecture was studied by KHINTCHINE [7], BESICOVITCH [1], SCHUR [18], and myself [2]. It was proved by MANN [12] in 1942. For the history of the problem, see the papers of ROHRBACH [16], ERDÖS and NIVEN [6], OSTMANN [15], and NIVEN [14].

The density of $A + B$ where $1 > \alpha > 0$ and $\beta = 0$ is not necessarily greater than α. KHINTCHINE [8] called a sequence B of density zero an essential component if for every sequence A of positive density $\alpha < 1$ the density of $A + B$ is greater than α. He showed that the sequence of the squares is an essential component. More exactly, KHINTCHINE proved:

If B is the sequence of the squares, then

$$\gamma \geq \alpha\{1 + \tfrac{1}{2} \cdot 10^{-8}(1 - \alpha)^2\}.$$

This result was improved by BUCHSTAB [4] and MORIMOTO [13].

An important generalization was obtained by ERDÖS [5]. A sequence of positive integers is called a basic sequence of order h if every positive integer is the sum of at most h elements of the sequence. ERDÖS showed that every basic sequence is an essential component. (The converse is false; see LINNIK [11].) More exactly, ERDÖS proved:

If A is of density α and if B is a basic sequence of order h, then

$$\gamma \geq \gamma' \geq \alpha\left(1 + \frac{1-\alpha}{2h}\right). \tag{3}$$

Assume that B contains 1 as element. For each m let $g(m)$ be the smallest number of elements of B whose sum equals m. Assume that the smallest upper bound λ of the quotients $\left\{\sum_{m=1}^{n} g(m)\right\}/n$ is finite. STÖHR [20] calls λ the mean order of B.

LANDAU [10] showed that in (3) the order h can be replaced by the mean order λ. Moreover, he proved that $\lambda \leq \frac{19}{6}$ for the sequence of the squares.

LANDAU's result was improved by me in 1938 [3] when I proved that

$$\gamma \geq \gamma' \geq \alpha\left(1 + \frac{1-\sqrt{\alpha}}{\lambda}\right) \tag{4}$$

and by S. SELBERG [19] in 1944 who proved that

$$\gamma \geq \gamma' \geq \alpha\left\{1 + \frac{3(1-\alpha)}{4\lambda}\right\}. \tag{5}$$

(5) is better than (4) for $\alpha > \frac{1}{9}$ while (4) is better for $\alpha < \frac{1}{9}$.

In this paper these results will be improved further for $\frac{1}{9} \leq \alpha \leq \frac{1}{2}$. *It will be shown that for* $\alpha \leq \frac{1}{2}$

$$\gamma \geq \gamma' \geq \alpha + \tfrac{1}{2}\left\{\lambda - \sqrt{\lambda^2 - 3(\alpha - \alpha^2)}\right\}. \tag{6}$$

It is easy to see that (6) is better than (5) for $\frac{1}{9} \leq \alpha \leq \frac{1}{2}$ and better than (4) for $\alpha = \frac{1}{9}$. The case $\alpha > \frac{1}{2}$ will be studied in another paper.

The following example shows that (6) is sometimes better than (2). Let A be the sequence of the numbers which are congruent to 1 (mod 3), B the sequence consisting of the squares and of the numbers which are congruent to 2 (mod 90). Then the density of B equals $\frac{1}{90}$ and its mean order is less than or equal to $\frac{19}{6}$. It follows from (2) that $\gamma - \frac{1}{3} \geq \frac{1}{90}$ and from (6) that

$$\gamma - \frac{1}{3} \geq \frac{1}{2}\left(\frac{19}{6} - \sqrt{\frac{361}{36} - 1 + \frac{1}{3}}\right) = \frac{19 - \sqrt{337}}{12} > \frac{1}{20}.$$

In the following, we denote the sequence of the integers which are not elements of A by $\bar{A}$. For abbreviation, we set $A(n) = N$.

LEMMA 1 [SELBERG 19]. *We have*

$$\sum_{k=1}^{n} A(k)\{A(k) - A(k-1)\} = \tfrac{1}{2}(N^2 + N).$$

PROOF. We have

$$A(k)\{A(k)-A(k-1)\}=A(k) \quad \text{if} \quad k\in A,$$
$$A(k)\{A(k)-A(k-1)\}=0 \quad \text{if} \quad k\in \bar{A}.$$

Denoting the i-th element of A by a_i we obtain

$$\sum_{k=1}^{n} A(k)\{A(k)-A(k-1)\}=A(a_1)+A(a_2)+\cdots+A(a_N)$$
$$=1+2+\cdots+N=\tfrac{1}{2}(N^2+N).$$

LEMMA 2 [SELBERG 19]. *Let A be a sequence of positive integers with density $\alpha\leq\frac{1}{2}$. Assume*[1]) *that n is an element of $\bar{A}$. We set $[n/2]=z$ and*

$$\bar{A}(n-z-1)=\varrho(1-\alpha)(n-z-1). \tag{7}$$

Let $D_m(n)$ be the number of elements $a_\varkappa$ of A for which $a_\varkappa\leq n-m$ and for which $a_\varkappa+m\in\bar{A}$. Then

$$2\sum_{m=1}^{z} D_m(n)\geq N-N^2+2\alpha N(n-z-1)+$$
$$+\alpha(2nz-z^2+\alpha z^2-\alpha n^2+2\alpha n-z-\alpha).$$

PROOF. An element k of A is counted in $\sum_{m=1}^{z} D_m(n)$ as often as there are numbers m such that $k+m\leq n$ and $k+m\in\bar{A}$. If $k\leq n-z-1$, then k must be counted $\{\bar{A}(k+z)-\bar{A}(k)\}$ times while an element k of A for which $n-z\leq k\leq n-1$ must be counted $\{\bar{A}(n)-\bar{A}(k)\}$ times. Hence

$$\begin{cases}\displaystyle\sum_{m=1}^{z} D_m(n)=\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}\{\bar{A}(k+z)-\bar{A}(k)\}+\\ \displaystyle+\sum_{k=n-z}^{n-1}\{A(k)-A(k-1)\}\{\bar{A}(n)-\bar{A}(k)\}.\end{cases} \tag{8}$$

Since $\bar{A}(k)=k-A(k)$, we have

$$\bar{A}(k+z)-\bar{A}(k)=z-A(k+z)+A(k),$$
$$\bar{A}(n)-\bar{A}(k)=n-k-A(n)+A(k)=n-k-N+A(k),$$

hence by (8)

$$\begin{cases}\displaystyle\sum_{m=1}^{z} D_m(n)=\sum_{k=1}^{n-1}\{A(k)-A(k-1)\}A(k)+z\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}-\\ \displaystyle-\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+(n-N)\sum_{k=n-z}^{n-1}\{A(k)-A(k-1)\}-\\ \displaystyle-\sum_{k=n-z}^{n-1}k\{A(k)-A(k-1)\}.\end{cases} \tag{9}$$

[1]) This assumption is not necessary; it is made only to simplify the proof.

Since for any positive integers c and d with $c<d$

$$\sum_{k=c}^{d}\{A(k)-A(k-1)\}=A(d)-A(c-1),$$

it follows from (9) and Lemma 1 that

$$(10)\quad\begin{cases}\sum\limits_{m=1}^{z}D_m(n)=\frac{1}{2}(N^2+N)+zA(n-z-1)-\\ -\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+(n-N)\{A(n-1)-A(n-z-1)\}-\\ -\sum\limits_{k=n-z}^{n-1}k\{A(k)-A(k-1)\}.\end{cases}$$

Now

$$\sum_{k=n-z}^{n-1}k\{A(k)-A(k-1)\}=-\sum_{k=n-z}^{n-1}A(k)+nA(n-1)-(n-z)A(n-z-1).$$

Substituting this in (10) and using $A(n-1)=A(n)=N$ we obtain

$$(11)\quad\begin{cases}\sum\limits_{m=1}^{z}D_m(n)=\frac{1}{2}(N-N^2)+NA(n-z-1)-\\ -\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+\sum\limits_{k=n-z}^{n-1}A(k).\end{cases}$$

On the other hand

$$\begin{aligned}&\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\\ &\quad=\sum_{k=1}^{n-z-1}A(k+z)\{k-\bar{A}(k)-(k-1)+\bar{A}(k-1)\}\\ &\quad=\sum_{k=1}^{n-z-1}A(k+z)-\sum_{k=1}^{n-z-1}A(k+z)\{\bar{A}(k)-\bar{A}(k-1)\}\\ &\quad\le\sum_{k=1}^{n-z-1}A(k+z)-\alpha\sum_{k=1}^{n-z-1}(k+z)\{\bar{A}(k)-\bar{A}(k-1)\},\end{aligned}$$

$$(12)\quad\begin{cases}\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\le\sum\limits_{k=z+1}^{n-1}A(k)-\\ -\alpha\sum\limits_{k=1}^{n-z-1}k\{\bar{A}(k)-\bar{A}(k-1)\}-\alpha z\bar{A}(n-z-1).\end{cases}$$

Since $A(k)\ge\alpha k$, we have

$$\bar{A}(k)=k-A(k)\le(1-\alpha)k,$$

$$k\ge\bar{A}(k)/(1-\alpha).$$

It follows from (12) that

$$\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\leq\sum_{k=z+1}^{n-1}A(k)-$$
$$-\frac{\alpha}{1-\alpha}\sum_{k=1}^{n-z-1}\bar{A}(k)\{\bar{A}(k)-\bar{A}(k-1)\}-\alpha z\bar{A}(n-z-1)$$

and from Lemma 1 applied to $\bar{A}$ and from (7) that

$$\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\leq\sum_{k=z+1}^{n-1}A(k)-$$
$$-\frac{\alpha}{2(1-\alpha)}\{\bar{A}^2(n-z-1)+\bar{A}(n-z-1)\}-\alpha z\bar{A}(n-z-1)$$
$$=\sum_{k=z+1}^{n-1}A(k)-\frac{\alpha}{2}\{\varrho^2(1-\alpha)(n-z-1)^2+\varrho(n-z-1)+$$
$$+2\varrho z(1-\alpha)(n-z-1)\}.$$

Substituting this in (11) we obtain

$$(13)\quad\begin{cases}2\sum_{m=1}^{z}D_m(n)\geq N-N^2+2N\{n-z-1-\varrho(1-\alpha)(n-z-1)\}+\\+2\sum_{k=n-z}^{n-1}A(k)-2\sum_{k=z+1}^{n-1}A(k)+\alpha\varrho(n-z-1)\times\\\times\{\varrho(1-\alpha)(n-z-1)+1+2(1-\alpha)z\}.\end{cases}$$

Since $z=[n/2]$, we have $n-1\leq 2z$, hence $n-z\leq z+1$, and by (13)

$$(14)\quad\begin{cases}2\sum_{m=1}^{z}D_m(n)\geq N-N^2+2N(n-z-1)\{1-\varrho(1-\alpha)\}+2\sum_{k=n-z}^{z}A(k)+\\+\alpha\varrho(n-z-1)\{\varrho(1-\alpha)(n-z-1)+1+2(1-\alpha)z\}.\end{cases}$$

Denoting the right-hand side of (14) by $G(\varrho)$ we have

$$(15)\quad\begin{cases}\frac{\partial G}{\partial\varrho}=-2N(n-z-1)(1-\alpha)+2\alpha(1-\alpha)(n-z-1)^2\varrho+\\+\alpha(n-z-1)+2\alpha(1-\alpha)z(n-z-1).\end{cases}$$

Since $\partial G/\partial\varrho$ is a linear function of ϱ and since $\partial^2 G/\partial\varrho^2$ is positive, it follows that $G(\varrho)$ has one minimum and no maximum. If we denote the value of ϱ for which $G(\varrho)$ takes its minimum by ϱ_0, then it follows from (15) that

$$(16)\qquad\varrho_0=\frac{2N(1-\alpha)-\alpha-2\alpha(1-\alpha)z}{2\alpha(1-\alpha)(n-z-1)}\geq\frac{2n(1-\alpha)-1-2(1-\alpha)z}{2(1-\alpha)(n-z-1)}.$$

Now $\alpha\leq\frac{1}{2}$, hence $1\leq 2(1-\alpha)$, and by (16)

$$(17)\qquad\varrho_0\geq\frac{2n-2-2z}{2(n-z-1)}=1.$$

On the other hand, it follows from (7) that we have to consider only the interval $0 \leq \varrho \leq 1$. By (17), the function $G(\varrho)$ takes its smallest value in the interval $0 \leq \varrho \leq 1$ for $\varrho = 1$. Therefore by (14) if we replace $A(k)$ by αk and add

$$2\sum_{m=1}^{z} D_m(n) \geq N - N^2 + 2\alpha N(n-z-1) + \alpha n(2z-n+1) +$$
$$+ \alpha(n-z-1)\{(1-\alpha)(n+z-1)+1\}$$
$$= N - N^2 + 2\alpha N(n-z-1) + 2\alpha n z - \alpha n^2 + \alpha n +$$
$$+ (\alpha - \alpha^2)(n^2 - z^2 - 2n + 1) + \alpha n - \alpha z - \alpha.$$
$$2\sum_{m=1}^{z} D_m(n) \geq N - N^2 + 2\alpha N(n-z-1) +$$
$$+ \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha).$$

LEMMA 3 [LANDAU *10*]. *Let A and B be sequences of positive integers and S the sequence of the numbers*

$$a + e_2 b \quad (a \in A,\ b \in B;\ e_2 = 0 \text{ or } 1).$$

Assume that B contains the number 1 as element. Let $g(m)$ be the smallest number of elements of B whose sum equals m. Assume that the mean order λ of B is finite. Using the notations of Lemma 2 we have

$$\sum_{m=1}^{z} D_m(n) \leq \lambda z\{S(n) - N\}.$$

PROOF. We write m as sum of $g(m)$ elements of B

$$m = b_1 + b_2 + \cdots + b_{g(m)}.$$

Assume that $a \in A$ while $a + m \in \bar{A}$ and $a + m \leq n$. We consider the numbers

$$(18) \qquad a,\ a + b_1,\ a + b_1 + b_2, \ldots,\ a + b_1 + b_2 + \cdots + b_{g(m)} = a + m.$$

The first of these numbers is an element of A while the last is an element of $\bar{A}$. Hence there exists a smallest number f such that

$$a + b_1 + b_2 + \cdots + b_{f-1} \in A$$

while

$$s = a + b_1 + b_2 + \cdots + b_{f-1} + b_f \in \bar{A}.$$

Since s is the sum of an element of A and the element b_f of B, it is an element of S. Hence to each element a which must be counted in $D_m(n)$ belongs a sequence (18), hence at least one element s of S which does not belong to A. The number of such elements $s \leq n$ is $S(n) - N$. Each such s can appear in at most $g(m)$ sequences (18) since it can be the second element, the third element, ..., the $\{g(m)+1\}$ st element of the sequence. Hence

$$(19) \qquad D_m(n) \leq g(m)\{S(n) - N\},$$

$$(20) \qquad \sum_{m=1}^{z} D_m(n) \leq \lambda z\{S(n) - N\}.$$

We now want to improve Lemma 3.

LEMMA 4. *Assume that* $n \in \bar{S}$. *We set* $z = [n/2]$. *Using the notations of Lemma 3 we have*

$$2\sum_{m=1}^{z} D_m(n) \leq (2\lambda z + 1)\{S(n) - N\} - \{S(n) - N\}^2.$$

PROOF. Let s_1 be an element of S which does not belong to A. This element will appear in exactly $g(m)$ sequences (18) if all the numbers

$$(21) \quad s_1 - b_1, \quad s_1 - b_1 - b_2, \quad s_1 - b_1 - b_2 - b_3, \quad \ldots, \quad s_1 - b_1 - b_2 - \cdots - b_{g(m)}$$

are elements of A. For certain values of s_1 and for certain values of m all the numbers (21) can belong to A. But the sequence (21) can contain elements of $\bar{A}$ or non-positive numbers.

In particular, let s_2 be another element of S which does not belong to A, and assume that

$$(22) \qquad 0 < s_1 - s_2 \leq z.$$

If we choose $m = s_1 - s_2$ and construct for s_1 and m the sequence (21), then the last element is s_2 which belongs to $\bar{A}$. Hence the sequence (21) contains at most $g(m) - 1$ elements of A, and s_1 appears in at most $g(m) - 1$ sequences (18) for $m = s_1 - s_2$.

Similarly, if $s \leq z$ is an element of S which does not belong to A, then for $m = s, s+1, \ldots, z$ the last element of the sequence (21) is not positive. Hence for such an m and s the element s appears in at most $g(m) - 1$ sequences (18).

For $1 \leq m \leq z$ we denote the number of pairs of elements s_1 and s_2 of S less than or equal to n which do not belong to A and for which $s_1 - s_2 = m$ by $\varphi(m)$ and the number of elements of S which do not belong to A and which are less than or equal to m by $\psi(m)$. Then it follows that there exist at least $\varphi(m) + \psi(m)$ elements of S for each m which do not belong to A and for which the last element of the sequence (21) is not in A. Hence we obtain instead of (19)

$$\begin{aligned} D_m(n) &\leq \{g(m) - 1\}\{\varphi(m) + \psi(m)\} + g(m)\{S(n) - N - \varphi(m) - \psi(m)\} \\ &= g(m)\{S(n) - N\} - \varphi(m) - \psi(m), \end{aligned}$$

and (20) can be replaced by

$$(23) \qquad \sum_{m=1}^{z} D_m(n) \leq \lambda z\{S(n) - N\} - \sum_{m=1}^{z}\{\varphi(m) + \psi(m)\}.$$

Let P_1 be the total number of pairs of elements of S less than or equal to n which do not belong to A and which satisfy (22), and P_2 be the number of such pairs for which

$$(24) \qquad s_1 - s_2 > z \geq \tfrac{1}{2}(n-1).$$

Then

$$(25) \qquad P_1 + P_2 = \tfrac{1}{2}\{S(n) - N\}^2 - \tfrac{1}{2}\{S(n) - N\}.$$

Moreover

$$P_1 = \sum_{m=1}^{z} \varphi(m). \tag{26}$$

We maintain that

$$P_2 \leq \sum_{m=1}^{z} \psi(m). \tag{27}$$

Since $n \in \overline{S}$, it follows that $s_1 \leq n-1$, hence by (24)

$$s_2 + z + 1 \leq s_1 \leq n - 1. \tag{28}$$

Therefore for a given s_2, there exist at most $n-1-s_2-z$ values s_1 satisfying (24). Moreover, by (28)

$$s_2 \leq n - 2 - z \leq 2z - 1 - z = z - 1.$$

It follows that

$$P_2 \leq \sum_{s_2 \leq z-1} (n - 1 - s_2 - z). \tag{29}$$

On the other hand, we have to count s_2 in $\sum_{m=1}^{z} \psi(m)$ for $m = s_2, s_2+1, \ldots, z$, hence $z - s_2 + 1$ times,

$$\sum_{m=1}^{z} \psi(m) = \sum_{s_2 \leq z} (z - s_2 + 1) \geq \sum_{s_2 \leq z-1} (z - s_2 + 1). \tag{30}$$

But

$$n - 1 - s_2 - z \leq 2z - s_2 - z < z - s_2 + 1,$$

hence

$$\sum_{s_2 \leq z-1} (n - 1 - s_2 - z) \leq \sum_{s_2 \leq z-1} (z - s_2 + 1), \tag{31}$$

and (27) follows from (29), (31), and (30).

Now we obtain from (26), (27), and (25)

$$\sum_{m=1}^{z} \{\varphi(m) + \psi(m)\} \geq \tfrac{1}{2}\{S(n) - N\}^2 - \tfrac{1}{2}\{S(n) - N\}$$

and from (23)

$$2\sum_{m=1}^{z} D_m(n) \leq (2\lambda z + 1)\{S(n) - N\} - \{S(n) - N\}^2.$$

LEMMA 5. *Using the same notations we have*

$$S(n) \geq N + \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 + N^2 - 2N\alpha(n - z - 1) - N - \\ - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha)\}^{\frac{1}{2}}. \tag{32}$$

PROOF. It follows from Lemma 4 that

$$S(n) - N \geq \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 - 2\sum_{m=1}^{z} D_m(n)\}^{\frac{1}{2}}.$$

Now (32) follows from Lemma 2.

LEMMA 6. *The function*

$$f(x) = x + u - (v - wx + x^2)^{\frac{1}{2}}$$

where $v > 0$, $w > 0$, *and* $w^2 < 4v$ *is monotonic increasing.*

PROOF. We have

$$f'(x) = 1 - (x - \tfrac{1}{2}w)(v - wx + x^2)^{-\frac{1}{2}}.$$

If $x \leq \frac{1}{2}w$, then $f'(x)$ is positive. Assume now that $x > \frac{1}{2}w$. In order to prove that $f'(x)$ is positive it is sufficient to show that

$$(x - \tfrac{1}{2}w) < (v - wx + x^2)^{\frac{1}{2}}.$$

But this is correct since $w^2 < 4v$.

LEMMA 7. The inequality (32) remains correct if N is replaced by αn.

PROOF. We want to prove that the right-hand side of (32) satisfies the assumptions of Lemma 6 for $x = N$, $u = \lambda z + \frac{1}{2}$,

$$v = (\lambda z + \tfrac{1}{2})^2 - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha), \tag{33}$$

$$w = 2\alpha(n - z - 1) + 1. \tag{34}$$

We distinguish between two cases.

I. *n is even.* Here $z = \frac{1}{2}n$. Since $\lambda > 1$, we have by (33)

$$4v > n^2 + 2n + 1 - 4\alpha n^2 + \alpha n^2 - \alpha^2 n^2 + 4\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n + 4\alpha^2,$$

$$4v > n^2 + 2n + 1 - 3\alpha n^2 + 3\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n + 4\alpha^2. \tag{35}$$

Moreover, by (34)

$$w^2 = \{\alpha(n - 2) + 1\}^2 = \alpha^2 n^2 - 4\alpha^2 n + 4\alpha^2 + 2\alpha n - 4\alpha + 1,$$

hence by (35)

$$4v - w^2 > n^2 + 2n - 3\alpha n^2 + 2\alpha^2 n^2 - 4\alpha^2 n + 4\alpha. \tag{36}$$

If we denote the right-hand side of (36) by $h(\alpha)$, then $\frac{\partial h(\alpha)}{\partial \alpha} = 0$ if

$$-3n^2 + 4\alpha n^2 - 8\alpha n + 4 = 0,$$

hence for $\alpha = \alpha_0$ where

$$\alpha_0 = \frac{3n^2 - 4}{4n^2 - 8n} > \frac{1}{2}$$

because $n > 2$. Since $\frac{\partial^2 h(\alpha)}{\partial \alpha^2}$ is positive, it follows that $h(\alpha)$ takes its smallest value in the interval $0 \leq \alpha \leq \frac{1}{2}$ for $\alpha = \frac{1}{2}$, hence by (36)

$$4v - w^2 > n^2 + 2n - \tfrac{3}{2}n^2 + \tfrac{1}{2}n^2 - n + 2 > 0,$$

hence $4v > w^2$ and $v > 0$.

II. *n is odd.* Here $z=\frac{1}{2}(n-1)$. It follows from (33) and (34) that

$$4v > n^2 - 4\alpha n^2 + 4\alpha n + \alpha n^2 - 2\alpha n + \alpha - \alpha^2 n^2 + 2\alpha^2 n - \\ - \alpha^2 + 4\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n - 2\alpha + 4\alpha^2$$
$$= n^2 - 3\alpha n^2 + 3\alpha^2 n^2 + 4\alpha n - 6\alpha^2 n - \alpha + 3\alpha^2,$$
$$w^2 = \{\alpha(n-1)+1\}^2 = \alpha^2 n^2 - 2\alpha^2 n + \alpha^2 + 2\alpha n - 2\alpha + 1,$$

$$4v - w^2 > n^2 - 3\alpha n^2 + 2\alpha^2 n^2 + 2\alpha n - 4\alpha^2 n + \alpha + 2\alpha^2 - 1. \tag{37}$$

Let α_0 be the value of α for which the right-hand side of (37) takes its minimum, then

$$-3n^2 + 4\alpha_0 n^2 + 2n - 8\alpha_0 n + 1 + 4\alpha_0 = 0,$$
$$\alpha_0 = \frac{3n^2 - 2n - 1}{4n^2 - 8n + 4} = \frac{3n+1}{4n-4} > \frac{1}{2}$$

since $n > 1$. It follows as above that the right-hand side of (37) takes its smallest value in the interval $0 \leq \alpha \leq \frac{1}{2}$ for $\alpha = \frac{1}{2}$, hence by (37)

$$4v - w^2 > n^2 - \tfrac{3}{2}n^2 + \tfrac{1}{2}n^2 + n - n + \tfrac{1}{2} + \tfrac{1}{2} - 1 = 0,$$

hence $4v > w^2$.

Therefore in both cases, the right-hand side of (32) satisfies the assumptions of Lemma 6. It follows from this Lemma that (32) remains correct if N is replaced by αn. We obtain

$$S(n) \geq \alpha n + \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 - \alpha n + \alpha^2 n^2 - 2\alpha^2 n(n - z - 1) - \\ - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha)\}^{\frac{1}{2}},$$

$$S(n) \geq \alpha n + \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 - \alpha z(1-\alpha)(2n - z) - \alpha(n-z) + \alpha^2\}^{\frac{1}{2}}. \tag{38}$$

LEMMA 8 [SELBERG *19*]. *If B is a sequence of density $\beta \leq \frac{1}{2}$, then the mean order λ of B is greater than or equal to $\frac{3}{2}$.*

PROOF. For the $B(n)$ numbers m which are elements of B and which are less than or equal to n, we have $g(m) = 1$. For the remaining $n - B(n)$ numbers m which are less than or equal to n we have $g(m) \geq 2$. Hence for $n = 1, 2, \ldots$ it follows that $\lambda n \geq B(n) + 2\{n - B(n)\}$,

$$B(n) \geq (2 - \lambda) n \qquad (n = 1, 2, \ldots). \tag{39}$$

On the other hand, if (39) holds, then the density β of B is

$$\beta \geq 2 - \lambda,$$

hence

$$\lambda \geq 2 - \beta \geq 2 - \tfrac{1}{2} = \tfrac{3}{2}.$$

LEMMA 9. *For $0 \leq \alpha \leq \frac{1}{2}$ we have*

$$\alpha + \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} < 1.$$

PROOF. If the Lemma is not correct, then

$$\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)} \geq 2 - 2\alpha, \tag{40}$$

$$\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} \geq 2(1-\alpha)\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\},$$

$$3\alpha(1-\alpha) \geq 2(1-\alpha)\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}$$

$$3\alpha \geq 2\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

This is impossible since $\lambda > 1$ and $\alpha \leq \frac{1}{2}$. Hence (40) is not true, and the Lemma is proved.

Using these lemmas we can prove

THEOREM. *If A is a sequence of positive integers of density $\alpha \leq \frac{1}{2}$, and B a sequence of mean order $\lambda \geq \frac{3}{2}$, then the density γ' of the sequence S of the integers of the form*

$$a + e_2 b \quad (a \in A, \quad b \in B; \quad e_2 = 0 \text{ or } 1)$$

satisfies the inequality

$$\gamma' \geq \alpha + \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}. \tag{41}$$

By Lemma 8, the condition $\lambda \geq \frac{3}{2}$ is satisfied if B is of density 0 or of positive density $\beta \leq \frac{1}{2}$.

PROOF. We want to prove that for $n = 1, 2, \ldots$

$$S(n) > \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}. \tag{42}$$

This is trivial for $n = 1$ and $n = 2$ since 1 is in A, hence $2 \in S$, and since the right-hand side of (42) is less than n by Lemma 9. We now assume that (42) is already proved for all numbers less than n where $n \geq 3$. Hence

$$S(n-1) > \alpha(n-1) + \tfrac{1}{2}(n-1)\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

If $n \in S$, then

$$S(n) = S(n-1) + 1 > \alpha(n-1) + \tfrac{1}{2}(n-1)\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} + 1$$
$$= \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} + 1 - \alpha - \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\},$$

hence by Lemma 9

$$S(n) > \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

Therefore it is sufficient to prove (42) for the case that $n \in \bar{S}$.

We set $z = [n/2]$. Then the assumptions of the Lemmas 4, 5, and 7 hold. We have to distinguish between two cases.

I. *n is even.* It follows from (38) that

$$S(n) \geq \alpha n + \tfrac{1}{2}(\lambda n + 1) - \tfrac{1}{2}\{(\lambda n + 1)^2 - 3\alpha(1-\alpha) n^2 - 2\alpha n + 4\alpha^2\}^{\frac{1}{2}},$$

$$S(n) \geq \alpha n + \tfrac{1}{2}(\lambda n + 1) - \tfrac{1}{2}\{n^2(\lambda^2 - 3\alpha + 3\alpha^2) + 2n(\lambda - \alpha) + 1 + 4\alpha^2\}^{\frac{1}{2}}. \tag{43}$$

Now (42) follows from (43) if we can prove that

$$(44)\quad 1-\{n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda-\alpha)+1+4\alpha^2\}^{\frac{1}{2}}>-\{n^2(\lambda^2-3\alpha+3\alpha^2)\}^{\frac{1}{2}}.$$

Hence we have to show that

$$1+n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda^2-3\alpha+3\alpha^2)^{\frac{1}{2}}$$
$$>n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda-\alpha)+1+4\alpha^2,$$
$$n(\lambda^2-3\alpha+3\alpha^2)^{\frac{1}{2}}>\lambda n-\alpha n+2\alpha^2,$$
$$\lambda^2n^2-3\alpha n^2+3\alpha^2n^2>\lambda^2n^2+\alpha^2n^2+4\alpha^4-2\alpha\lambda n^2+4\alpha^2\lambda n-4\alpha^3 n,$$
$$n^2(2\alpha\lambda-3\alpha+2\alpha^2)-4\alpha^2\lambda n+4\alpha^3 n-4\alpha^4>0.$$

Since $4\alpha^3 n-4\alpha^4>0$, it is sufficient to show that

$$n(2\lambda-3+2\alpha)-4\alpha\lambda\geq 0.$$

The coefficient of n is greater than 0, hence it is sufficient that

$$(45)\qquad n\geq\frac{4\alpha\lambda}{2\lambda-3+2\alpha}=\frac{4\alpha}{2-(3-2\alpha)/\lambda}.$$

The right-hand side is a decreasing function of λ. Since $\lambda\geq\frac{3}{2}$, it follows from (45) that $n\geq 3$ is sufficient. But this is true. Therefore (44) is proved, and (42) holds if n is even.

II. *n is odd.* Here $z=\frac{1}{2}(n-1)$. It follows from (38) that

$$S(n)\geq\alpha n+\tfrac{1}{2}(\lambda n-\lambda+1)-\tfrac{1}{2}\{(\lambda n-\lambda+1)^2-$$
$$-\alpha(n-1)(1-\alpha)(3n+1)-2\alpha(n+1)+4\alpha^2\}^{\frac{1}{2}},$$

$$(46)\quad\begin{cases}S(n)\geq\alpha n+\tfrac{1}{2}(\lambda n-\lambda+1)-\tfrac{1}{2}\{(\lambda n-\lambda+1)^2-3\alpha n^2+\\ \qquad+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}.\end{cases}$$

We maintain that

$$(47)\quad\begin{cases}-\lambda+1-\{(\lambda n-\lambda+1)^2-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}\\ \qquad>-(\lambda^2n^2-3\alpha n^2+3\alpha^2n^2)^{\frac{1}{2}}.\end{cases}$$

Hence we have to prove that

$$-\lambda+1+\{\lambda^2n^2-3\alpha n^2+3\alpha^2n^2\}^{\frac{1}{2}}$$
$$>\{(\lambda n-\lambda+1)^2-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}.$$

The left-hand side is positive since

$$n^2(\lambda^2-3\alpha+3\alpha^2)\geq n^2(\lambda^2-\tfrac{3}{4})>\lambda^2-1>(\lambda-1)^2.$$

Hence it is sufficient to prove that

$$\lambda^2-2\lambda+1+\lambda^2n^2-3\alpha n^2+3\alpha^2n^2-2(\lambda-1)(\lambda^2n^2-3\alpha n^2+3\alpha^2n^2)^{\frac{1}{2}}$$
$$>\lambda^2n^2+\lambda^2+1-2\lambda^2n+2\lambda n-2\lambda-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2,$$

$$(48)\qquad 2\lambda^2n-2\lambda n+2\alpha^2n+\alpha-3\alpha^2>2(\lambda-1)n\{\lambda^2-3\alpha+3\alpha^2\}^{\frac{1}{2}}.$$

The left-hand side is greater than $2\lambda^2 n - 2\lambda n$ since $n \geq 3$ while the right-hand side is less than $2\lambda^2 n - 2\lambda n$. Hence (48) and (47) are true, and (42) follows from (46) when n is odd.

Bibliography

[1] Besicovitch, A.: On the density of the sum of two sequences of integers. J. London Math. Soc. **10**, 246—248 (1935). — [2] Brauer, Alfred: Über die Dichte der Summe von Mengen positiver ganzer Zahlen. I. Ann. of Math. **39**, 322—340 (1938). — II. Ann. of Math. **42**, 959—988 (1941). — [3] Brauer, Alfred: Über die Dichte der Summe zweier Mengen, deren eine von positiver Dichte ist. Math. Z. **44**, 212—232 (1938). — [4] Buchstab, A.: Über ein metrisches Problem der additiven Zahlentheorie. Matematitscheskij Sbornik **40**, 190—195 (1933). — [5] Erdös, P.: On the arithmetical density of the sum of two sequences one of which forms a basis for the integers. Acta arithmetica **1**, 197—200 (1936). — [6] Erdös, P., and I. Niven: The $\alpha + \beta$ hypothesis and related problems. Amer. Math. Monthly **53**, 314—317 (1946). — [7] Khintchine, A.: Zur additiven Zahlentheorie. Matematitscheskij Sbornik **39**, 27—34 (1932). — [8] Khintchine, A.: Über ein metrisches Problem der additiven Zahlentheorie. Matematitscheskij Sbornik **40**, 180—189 (1933). — [9] Landau, E.: Die Goldbachsche Vermutung und der Schnirelmannsche Satz. Göttinger Nachrichten, Math.-phys. Kl. **1930**, 255—276. — [10] Landau, E.: Über einige neuere Fortschritte der additiven Zahlentheorie, S. 3—6 and 60—63. Cambridge 1937. — [11] Linnik, U. V.: On Erdös's theorem on the addition of numerical sequences. Matematitscheskij Sbornik, N. S. **10**, 67—78 (1942). — [12] Mann, H. B.: A proof of the fundamental theorem on the density of sums of sets of positive integers. Ann. of Math. (2) **43**, 523—527 (1942). — [13] Morimoto, S.: Über einen Satz von Khintchine für eine Summenfolge. Tôhoku Math. J. **43** 1—3, (1937). — [14] Niven, I.: The asymptotic density of integers. Bull. Amer. Math. Soc. **57**, 420—434 (1951). — [15] Ostmann, H.-H.: Über die Dichten additiv komponierter Zahlenmengen. Arch. Math. **1**, 393—401 (1848/49). [16] Rohrbach, H.: Einige neuere Untersuchungen über die Dichte in der additiven Zahlentheorie. Jber. dtsch. Math.-Ver. **48**, 199—236 (1938). — [17] Schnirelmann, L.: Über additive Eigenschaften von Zahlen. Math. Ann. **107**, 649—690 (1933). — [18] Schur, I.: Über den Begriff der Dichte in der additiven Zahlentheorie. Sitzgsber. preuß. Akad. Wiss. Phys.-math. Kl. **1936**, 269—297. — [19] Selberg, Sigmund: Note on a metrical problem in the additive theory of numbers. Arch. Math. Naturvidensk. **47**, 111—118 (1944). — [20] Stöhr, A.: Bemerkungen zur additiven Zahlentheorie. I. Mittlere Ordnung. J. reine u. angew. Math. **183**, 168—174 (1941).

University of North Carolina

(Eingegangen am 28. Oktober 1955)